

UNIVERSIDADE FEDERAL DO PARANÁ
SUZANA SALETE DE SOUZA

**ANÁLISE DA INFORMAÇÃO JURÍDICA: INFOMETRIA APLICADA À
JUSRISPRUDÊNCIA SOBRE O PREGÃO**

CURITIBA

2014

SUZANA SALETE DE SOUZA

**ANÁLISE DA INFORMAÇÃO JURÍRICA: INFOMETRIA APLICADA À
JURISPRUDÊNCIA SOBRE O PREGÃO**

Trabalho de Conclusão de Curso apresentado como requisito parcial à obtenção do título de Bacharel em Gestão da Informação, do Curso de Gestão da Informação, da Universidade Federal do Paraná.

Orientador: Prof. Edson R. Guarido Filho

CURITIBA

2014

AGRADECIMENTOS

Agradeço primeiramente a Deus, por tudo o que tem feito e tudo que ainda vai fazer.

Aos meus pais, Lenir e Valdir, e minha irmã Rosana pelo apoio e ajuda sempre que precisei.

Ao meu marido Paulo Guilherme, que me ajudou e compreendeu nos momentos em que me ausentei por motivos acadêmicos, pela motivação, pelo amor e paciência e por estar sempre ao meu lado.

Aos amigos que conquistei durante estes anos, especialmente a Helisane, Valquíria e Bianca, que estiveram comigo desde o princípio me ajudando, me dando forças, e principalmente pela amizade que construímos e pelas boas lembranças que levarei.

Ao meu orientador, Prof. Edson Guarido, pela dedicação e paciência e a todos os meus professores que me proporcionaram tamanho conhecimento e crescimento pessoal.

Enfim, agradeço a todos que, de alguma forma, contribuíram para a concretização desta grande etapa em minha vida.

RESUMO

O presente estudo tem como tema a análise da informação jurídica, especificamente na jurisprudência, a partir de técnicas infométricas. O problema de pesquisa visa criar formas de tratar e visualizar a informação jurídica. Para responder ao problema o objetivo geral foi desenvolver um estudo a fim de analisar essa informação a partir de técnicas infométricas e verificar sua possível utilidade na recuperação da informação e análise da informação. O levantamento de dados consistiu na busca de acórdãos e súmulas no site do Tribunal de Contas da União (TCU). A análise das informações obtidas ocorreu utilizando técnicas de nuvem de palavras, redes de ocorrência, fórmula do ponto de Transição de Goffman e manipulação dos dados em planilhas eletrônicas. Os resultados obtidos demonstram que é possível a análise e tratamento desses dados jurídicos por métodos quantitativos, de modo a permitir a melhor recuperação e visualização da informação.

Palavras-Chave: Infometria. Informação Jurídica. Jurisprudência. Visualização e tratamento da informação.

ABSTRACT

This study has covered the analysis of legal information, specifically in the case law, from infometrics techniques. The research problem aims to create ways to treat and view legal information. To address the problem the overall goal was to develop a study to analyze information from infometrics techniques and verify their potential use on information retrieval and information analysis. The survey consisted in seeking judgments in the Court of Audit (TCU) site. The analysis of legal information was performed by techniques as word cloud, network analysis, Goffman's transition point and counting in specific spreadsheets. The results obtained demonstrate that it is possible to analyze and processing of these legal data for quantitative methods, so as to provide a better retrieval and display of information.

Keywords: Infometrics. Legal Information. Jurisprudence. Viewing and processing information.

LISTA DE QUADROS

QUADRO 1 – TIPOLOGIA PARA A DEFINIÇÃO E CLASSIFICAÇÃO DOS TERMOS.....	14
---	----

LISTA DE TABELAS

TABELA 1– RESULTADO DE PALAVRAS AC_1329_2006.....	43
TABELA 2– RESULTADO DE PALAVRAS AC_817_2005.....	44
TABELA 3– RESULTADO DE PALAVRAS AC_2482_2007.....	44
TABELA 4– RESULTADO DE PALAVRAS ACÓRDÃOS JUNTOS.....	45
TABELA 5– RESULTADO CAMPO EMENTA PESQUISA POR ASSUNTO 1.....	46
TABELA 6– RESULTADO CAMPO EMENTA PESQUISA POR ASSUNTO 2.....	46
TABELA 7– RESULTADO CAMPO EMENTA PESQUISA POR ASSUNTO 3.....	47
TABELA 8– RESULTADO CAMPO EMENTA PESQUISA POR ASSUNTO 4.....	47
TABELA 9– RESULTADO CAMPO ACÓRDÃO PESQUISA POR ASSUNTO 1.....	48
TABELA 10– RESULTADO CAMPO ACÓRDÃO PESQUISA POR ASSUNTO 2.....	48
TABELA 11– RESULTADO CAMPO EMENTA MAIS ACÓRDÃO PESQUISA POR ASSUNTO 1.....	49
TABELA 12– RESULTADO CAMPO EMENTA MAIS ACÓRDÃO PESQUISA POR ASSUNTO 2.....	49
TABELA 13– RESULTADO CAMPO EMENTA PESQUISA LIVRE 1.....	50
TABELA 14– RESULTADO CAMPO EMENTA PESQUISA LIVRE 2.....	50
TABELA 15– RESULTADO CAMPO EMENTA PESQUISA ESPECÍFICA 1.....	51
TABELA 16– RESULTADO CAMPO EMENTA PESQUISA ESPECÍFICA 2.....	52

LISTA DE IMAGENS

IMAGEM 1–PÁGINA INICIAL DO SITE DO TRIBUNAL DE CONTAS DA UNIÃO.....	22
IMAGEM 2– PÁGINA DE JURISPRUDÊNCIA.....	23
IMAGEM 3– BUSCA PESQUISA ASSUNTO.....	24
IMAGEM 4– BUSCA POR COLUNAS.....	26
IMAGEM 5– RESULTADOS DE LINHAS: PESQUISA POR ASSUNTO.....	27
IMAGEM 6– MENSAGEM DE DESATIVAÇÃO DE SOLUÇÃO DE BUSCA.....	29

IMAGEM 7– PARÂMETROS DE PESQUISA.....	29
IMAGEM 8– PESQUISA AVANÇADA POR ARGUMENTOS.....	30
IMAGEM 9– MENSAGEM DE DESATIVAÇÃO PESQUISA POR CAMPOS.....	32
IMAGEM 10– PESQUISA EM ÁRVORE.....	32
IMAGEM 11– VISUALIZAÇÃO DO SISTEMA DE BUSCA EM ÁRVORE: PRIMEIRA DIVISÃO.....	33
IMAGEM 12– VISUALIZAÇÃO DO SISTEMA DE BUSCA EM ÁRVORE: SEGUNDA DIVISÃO.....	34
IMAGEM 13– APRESENTAÇÃO DOS ASSUNTOS DE RESULTADOS OBTIDOS.....	35
IMAGEM 14– RESULTADO ACÓRDÃO 1329_2006 – PLENÁRIO.....	53
IMAGEM 15– RESULTADO ACÓRDÃO 185_2005 – PRIMEIRA CÂMARA.....	53
IMAGEM 16– RESULTADO AC_2482_2007 – PLENÁRIO.....	54
IMAGEM 17– RESULTADO DOS ACÓRDÃOS REUNIDOS.....	55
IMAGEM 18– ANÁLISE NA EMENTA E ACÓRDÃO A PARTIR DO PONTO T DE GOFFMAN.....	55
IMAGEM 19– ANÁLISE NA EMENTA E ACÓRDÃO A PARTIR DA FÓRMULA ALTERNATIVA.....	56
IMAGEM 20– ANÁLISE NA EMENTA A PARTIR DO PONTO T DE GOFFMAN E SEM STOPWORDS.....	57
IMAGEM 21– ANÁLISE NA EMENTA A PARTIR DA FÓRMULA ALTERNATIVA E SEM STOPWORDS.....	58
IMAGEM 22– ANÁLISE NA EMENTA A PARTIR DO PONTO T DE GOFFMAN E COM STOPWORDS.....	58
IMAGEM 23–ANÁLISE NA EMENTA A PARTIR DA FÓRMULA ALTERNATIVA E COM STOPWORDS.....	59
IMAGEM 24– ANÁLISE NO ACÓRDÃO A PARTIR DO PONTO T DE GOFFMAN.....	59
IMAGEM 25–ANÁLISE NO ACÓRDÃO A PARTIR DA FÓRMULA ALTERNATIVA.....	60
IMAGEM 26– ANÁLISE DAS EMENTAS (PESQUISA LIVRE) A PARTIR DO PONTO T DE GOFFMAN.....	60
IMAGEM 27– ANÁLISE DAS EMENTAS (PESQUISA LIVRE) A PARTIR DA FÓRMULAALTERNATIVA DE GOFFMAN.....	61
IMAGEM 28– NUVEM DE PALAVRAS PESQUISA ESPECÍFICA.....	62
IMAGEM 29– NUVEM DE PALAVRAS PESQUISA ESPECÍFICA 2.....	63
IMAGEM 30– REDE DE LEIS.....	68

IMAGEM 31– ACÓRDÃOS CITADOS.....	69
IMAGEM 32– TODOS OS DOCUMENTOS CITADO	70
IMAGEM 33– REDE DE COOCORRENCIA ACÓRDÃOS.....	71

LISTA DE GRÁFICOS

GRÁFICO 1 – TOTAL DE DECISÕES POR MINISTRO.....	64
GRÁFICO 2 – DECISÕES DE 2003 A 2014.....	65
GRÁFICO 3 – TIPOS DOCUMENTAIS MAIS CITADOS.....	66
GRÁFICO 4 – CITAÇÕES POR ANO.....	67

SUMÁRIO

1 INTRODUÇÃO.....	9
1.1	O
PROBLEMA.....	10
1.2 OBJETIVOS GERAL E ESPECÍFICOS.....	10
1.3 JUSTIFICATIVA.....	11
2 FUNDAMENTAÇÃO TEÓRICA.....	12
2.1 AS METRIAS DA INFORMAÇÃO.....	12
2.2 REGISTRO INFORMACIONAL E METADADOS.....	16
2.3 REGISTRO JURÍDICO E JURISPRUDÊNCIA.....	16
3 PROCEDIMENTOS METODOLÓGICOS.....	18
3.1 DESCRIÇÃO.....	18
4 ANÁLISE DOS DADOS OBTIDOS.....	22
4.1 ANÁLISE DE DADOS DE SITES JURÍDICOS.....	22
4.1.1 TRIBUNAL DE CONTAS DA UNIÃO.....	22
4.2 ACÓRDÃO.....	36
4.3 APLICAÇÃO DE TÉCNICAS INFOMETRICAS SELECIONADAS.....	41
4.4 ANÁLISE DA APLICAÇÃO DAS TÉCNICAS.....	43
4.4.1 ANÁLISE DE PALAVRAS.....	43
4.4.1.1 ANÁLISE ISOLADA DE ACÓRDÃOS	43
4.4.1.2 ANÁLISE DE PALAVRAS PESQUISA POR ASSUNTO.....	45
4.4.1.3 ANÁLISE DE PALAVRAS PESQUISA LIVRE.....	49
4.4.1.4 ANÁLISE DE PALAVRAS PESQUISA ESPECÍFICA POR ASSUNTO.....	51
4.4.2 REPRESENTAÇÃO EM NUVEM.....	52
4.4.2.1 ANÁLISE ISOLADA DE ACÓRDÃOS EM NUVEM.....	52
4.4.2.2 ANÁLISE EM NUVEM DE ACÓRDÃOS PESQUISA POR ASSUNTO.....	55
4.4.2.3 ANÁLISE EM NUVEM DE PALAVRAS PESQUISA LIVRE..	61
4.4.2.4 ANÁLISE EM NUVEM DE PALAVRAS PESQUISA ESPECÍFICA.....	62
4.4.3 ANÁLISE DE AUTORIA.....	64
4.4.4 ANÁLISE DA DECISÃO.....	65
4.4.5 ANÁLISE DA CITAÇÃO – GENEALOGIA.....	66

4.4.6	DISCUSSÃO SOBRE POSSIBILIDADES DE APLICAÇÃO.....	71
4.5	SÍNTESE DA SEÇÃO.....	72
5	CONSIDERAÇÕES FINAIS.....	74
5.1	LIMITAÇÕES E SUGESTÕES PARA PESQUISAS FUTURAS.....	75
5.2	COMENTÁRIOS FINAIS.....	75
	REFERÊNCIAS.....	76
7		
	APÊNDICE A – TABELA DE DECISÕES DE CADA MINISTRO.....	79
	APÊNDICE B – TABELA DE CONTAGEM DE NÚMERO DE VEZES QUE A PALAVRA “PREGÃO” APARECE EM CADA DOCUMENTO (PESQUISA POR ASSUNTO).....	81
2		
	APÊNDICE C – TABELA DE CONTAGEM DE NÚMERO DE VEZES QUE A PALAVRA “PREGÃO” APARECE EM CADA DOCUMENTO (PESQUISA LIVRE).....	84
5		

1 INTRODUÇÃO

Não raramente, dados de diversas áreas do conhecimento, são objetos de estudos quantitativos, sendo tabulados e manipulados para fins de tomadas de decisão, observação de cenários e insights sobre o futuro. Estudos como esses não se observam com muita frequência na área jurídica.

O presente trabalho traz alguns conceitos e aplicabilidades da infometria e outras técnicas quantitativas e estatísticas de medição para avaliação da informação jurídica.

Aborda assim a relação desses estudos e sua importância como um conjunto de métodos para medir a informação. Neste sentido apresenta-os como uma possível alternativa para tratar e visualizar a informação jurídica e ajudar profissionais da área em suas buscas diárias de informação e seu entendimento sobre a mesma.

O campo de estudo neste trabalho restringe-se a informações obtidas junto a sites jurídicos que apresentam informações primárias. Em específico, considerou-se o site do Tribunal de Contas da União para a recuperação de material correspondente à jurisprudência, como acórdãos e súmulas. Procedeu-se a análise do referido site, avaliando sua atual forma de busca e recuperação de dados e, a partir daí, foram feitas aplicações de métodos infométricos em registros jurídicos selecionados. Assim ao final deste estudo pretende-se mostrar que é possível trabalhar com a informação jurídica de forma quantitativa, a fim de melhorar sua visualização e entendimento.

Na sequência, serão apresentados o problema, justificativa e objetivos do presente trabalho. Posteriormente, segue a revisão de literatura - com o propósito de conhecer os fundamentos teóricos e elementos específicos, que caracterizam-se objeto de estudo deste trabalho. Só então, serão detalhados a metodologia, considerando as formas de obtenção e análise de dados realizados durante a pesquisa. Mais adiante, são expostos os resultados obtidos para que, depois, sejam apresentadas as considerações finais e recomendações para futuros trabalhos.

1.1 O PROBLEMA

A área jurídica e seus respectivos sites possuem hoje, uma gama enorme de informações que são muito requeridas, principalmente, por advogados e profissionais que trabalham na área. No entanto caracterizam-se de difícil recuperação.

Isso por que a informação muitas vezes esta dispersa e mal organizada. Não raramente esses profissionais ou qualquer tipo de usuário passam horas tentando encontrar uma informação. Pode-se presumir a partir disso, que haja uma falha na indexação dos dados ou que a informação não possui um processo de tratamento adequado, antes de chegar até o usuário.

Assim sendo, o problema de pesquisa é verificar se existe a possibilidade de tratar dados jurídicos, a partir de aplicações de técnicas infométricas, de forma que garanta uma melhor visualização desta informação e conseqüentemente melhor entendimento dos dados recuperados.

1.2 OBJETIVOS GERAL E ESPECÍFICOS

O objetivo geral, que norteia esta pesquisa é desenvolver, a partir de técnicas infométricas, um estudo que ajude a tratar e visualizar a informação jurídica. Para que seja atingido este objetivo faz-se necessário que os seguintes objetivos específicos sejam atendidos:

- Verificar os procedimentos de busca de jurisprudência num site jurídico, representado no presente estudo pelo site do Tribunal de Contas da União;
- Medir registros jurídicos enquanto unidade informacional;
- Medir registros jurídicos selecionados a partir de técnicas infométricas;
- Verificar a possibilidade de extensão dos resultados obtidos em processos de busca e recuperação da informação.

1.3 JUSTIFICATIVA

O trabalho de profissionais da área jurídica exige estudos e pesquisas constantes, a fim de verificar as legislações e decisões tomadas. Neste processo rotineiro, uma importante ferramenta de auxílio para esses profissionais são os sites do Tribunal de Contas e Tribunais Federais. No entanto, estes sites apresentam um nível de dificuldade de busca e recuperação da informação muito grande e na maior parte das vezes, o profissional perde muito tempo em suas pesquisas, devido a um sistema ineficiente.

A infometria propõe-se a estudar todos os tipos de registros informacionais, no entanto é pouco aplicada na área jurídica, sendo que sua contribuição poderia ser de relativamente significativa para medir o uso da informação.

Miranda, D'Amore e Pinto (2013, p. 99) explicam bem esse processo quando dizem que “a recuperação inadequada da Legislação e da Jurisprudência tanto provoca insatisfação em seus usuários como pode gerar danos, especialmente aos juristas, englobando todos os que se utilizam dela.” Nesta linha de pensamento, voltamos a afirmar o quão difícil pode ser o trabalho desses profissionais que precisam destas informações para trabalhar e desenvolver suas atividades.

Desta forma o objetivo deste estudo é medir registros jurídicos com base em técnicas infométricas, a fim de melhor visualizar e tratar esses registros, demonstrando que a infometria também pode ajudar esta área de conhecimento. Seja na forma de facilitar o trato com a informação recuperada, seja na análise.

Além disso, pretende-se mostrar também que a informetria não se restringe a uma área de conhecimento, e sim que, ela pode ser aplicada na área jurídica sem qualquer prejuízo. Defende-se que é possível aproximar essas duas áreas que trabalham em caminhos diferentes. Assim pode-se afirmar que esse estudo é importante não só para a área de direito, que terá uma possível melhoria, devido a um primeiro passo para melhor tratamento de suas informações, mas também para expandir o campo de atuação da infometria e seu potencial de aplicação.

2 FUNDAMENTAÇÃO TEÓRICA

A contextualização ao tema abordado fazem-se necessárias para o entendimento das técnicas infométricas e aspectos que englobam esse trabalho.

2.1 AS METRIAS DA INFORMAÇÃO

As metrias da informação não são métodos novos. Segundo Santos e Kobashi (2009) embora as primeiras iniciativas tenham surgido no século XIX, é em meados do século XX que os estudos métricos da informação ganham força e legitimidade.

No século XX se expandiu e foi aí que surgiram métodos e leis muito importantes na área da ciência da informação. Como mostra a afirmação de Santin (2010, p. 109) “[...] Leis de Lotka (1926) e Price (1965), o estudo do núcleo e da dispersão de artigos em periódicos científicos, orientado pela Lei de Bradford (1934); e o estudo da frequência de termos em textos científicos, orientado pela Lei de Zipf (1935)[...]”.

Nesta época também aconteceram as primeiras discussões sobre a bibliometria. Nos estudos métricos, segundo Santin (2010, p. 109), as metodologias mais relevantes são a bibliometria, a cientometria, a infometria e a webometria. Por este motivo faremos uma breve recuperação de algumas teorias sobre as metrias citadas acima, a fim de entender seus conceitos e aplicações.

Como já foi dito os estudos da bibliometria não são recentes, Otle (1986, apud BUFREM; PRATES, 2005, p. 11) a definiu em 1934 pela primeira vez como a parte da bibliografia “que se ocupa da medida ou da quantidade aplicada ao livro”. Mais tarde seu conceito foi discutido por vários estudiosos e ainda permanece em constante evolução e discussão. Chapula (1998, p. 134) define a bibliometria como sendo “o estudo dos aspectos quantitativos da produção, disseminação e uso da informação registrada”. Neste sentido entendemos que a bibliometria tem como objeto de estudo artigos, livros, revistas, autores e usuários.

Já a cientometria, segundo Spinak (1996) é considerada como o método que analisa a dinâmica da ciência e tem como objeto disciplinas, campos científicos e

tecnológicos, além de artigos, patentes, teses e dissertações. Neste mesmo caminho Bufrem e Prates (2005, p.13) afirmam que “quando os métodos quantitativos são utilizados para estudar as atividades científicas ou técnicas do ponto de vista de sua produção ou comunicação costuma-se denominá-los cientometria, a ciência da ciência”. Podemos perceber que apesar de semelhantes, o objeto de estudo é diferente, deste modo prosseguiremos então para a infometria.

Infometria é um método quantitativo usado para medir registros informacionais, muitos autores consideram a infometria como um “guarda-chuva” da bibliometria e cientometria. Santos e Kobashi (2009, p. 159) definem desta forma:

A bibliometria tem como objetos de estudo os livros ou as revistas científicas, cujas análises se vinculam à gestão de bibliotecas e bases de dados. A cientometria preocupa-se com a dinâmica da ciência, como atividade social, tendo como objetos de análise a produção, a circulação e o consumo da produção científica. A infometria, por sua vez, abarca as duas primeiras, tendo desenvolvido métodos e ferramentas para mensurar e analisar os aspectos cognitivos da ciência. (SANTOS; KOBASHI, 2009, p. 159)

Chapula (1998, p.135) apresenta uma visão bem semelhante afirmando que:

“infometria é o estudo dos aspectos quantitativos da informação em qualquer formato e não apenas registros catalográficos ou bibliografias, referentes a qualquer grupo social, e não apenas aos cientistas. A infometria pode incorporar, utilizar e ampliar os muitos estudos de avaliação da informação que estão fora dos limites da bibliometria e cientometria”. (CHAPULA, 1998, p.135)

Assim percebe-se que o método pode ser aplicado em qualquer tipo de registro de informação, inclusive em registros de informações jurídicas, como por exemplo, um acordo, que mais tarde será o objeto deste estudo.

Por fim a webometria apresenta-se como o método para medir aspectos relacionados a sites e ambientes web. De acordo com Abraham (1997, apud SANTIN, 2010, p. 111), a webometria busca elucidar a relação entre os principais elementos da web, seus nós (domínios, sítios e páginas), suas conexões (nexos, ou links) e a matriz de ligações resultantes que envolvem toda a extensão da rede. São vários os fatores que podem ser estudados dentro desta proposta maior, como por exemplo, conteúdo, links, luminosidade e outros fatores de sites e/ou sítios web.

Um exemplo de estudo na área foi o de Vereeland (2000, apud SHINTAKU; ROBREDO; BAPTISTA, 2011, p. 318) que levantou questões sobre a visibilidade e luminosidades dos sítios das bibliotecas relacionadas com o ensino de

Direito, para uma classificação. Observou que a luminosidade dos sítios estudados seguia o padrão 80/20, como na Lei de Bradford, sem, no entanto encontrar relação entre a visibilidade e a luminosidade. Mas como esse, podemos ainda estudar vários outros fatores, como, por exemplo, o estudo de Rousseau (1997, apud SHINTAKU; ROBREDO; BAPTISTA, 2011, p. 318) sobre inlinks, que fez relações entre os domínios mais acessados e a distribuição de Lotka, com estimativa de selflink de 30%.

Para complementar Bufrem e Prates (2005, p. 16) apresentam um quadro que ilustra os conceitos estudados até o momento. É possível visualiza-lo abaixo no quadro 1:

Quadro 1 - Tipologia para definição e classificação dos termos

Tipologia para definição e classificação dos termos

Tipologia/ Subcampo	Bibliometria	Cientometria	Informetria/Infometria	Webometria
Objeto de estudo	Livros, documentos, revistas, artigos, autores, usuários.	Disciplinas, assuntos, campos científicos e tecnológicos, patentes, dissertações e teses.	Palavras, documentos, banco de dados, comunicações informais (inclusive em âmbito não científico) e <i>homepage</i> na WWW.	Sítios na WWW, (URL, título, tipo, domínio, tamanho e <i>links</i>) motores de busca.
Variáveis	Número de empréstimos (circulação) e de citações, frequência de extensão de frases.	Fatores que se diferenciam as subdisciplinas. Como os cientistas se comunicam.	Medir a recuperação, relevância e revocação	Número de páginas por eixo, número de linhas por eixo, número de <i>links</i> que remetem ao mesmo sítio, "situações", estratégias de busca
Métodos	<i>Ranking</i> , frequência, distribuição.	Análise de conjunto de correspondência, co-ocorrência de termos, expressões, palavras-chave.	Modelo vetor espaço, modelos booleanos de recuperação, modelos probabilísticos, linguagem de processamento, abordagens baseada no conhecimento, tesouros.	Fator de Impacto da Web (FIW), densidade dos <i>links</i> , "situações", estratégias de busca
Objetivos	Alocar recursos, pessoas, tempo, dinheiro.	Identificar domínios de interesse, compreender como e quanto os cientistas se comunicam.	Melhorar a eficiência da recuperação da informação, identificar relações entre os diversos sistemas de informação.	Avaliar o sucesso de determinados sítios, detectar a presença de instituições, pesquisadores na rede e melhorar a eficiência dos motores de busca na recuperação das informações.

Fonte: Adaptado de Bufrem e Prates (2005, p.16).

Neste as autoras, acrescentam a contribuição da Webometria ao quadro enfatizando que seu objetivo é de avaliar o sucesso de determinados sítios, destacar a presença de instituições e pesquisadores e melhorar a eficiência dos motores de busca. O que se assemelha com o objetivo deste trabalho novamente.

Assim é importante ressaltar que apesar de caracterizarem estudos com objetos diferentes, muitas vezes esses métodos podem se relacionar. Neste trabalho, por

exemplo, utilizaremos a infometria como método de estudo por abranger qualquer registro informacional, no entanto em alguns momentos também usaremos a webometria, para análise dos sites e ambiente web. Esses métodos não trabalham isolados, e sim quase sempre em conjunto, um completando o outro.

Para complementar o estudo, é importante também conhecer conceitos e aplicações de leis conhecidas nesta área, como a lei de Lotka, lei de Bradford e lei de Zipf, esta última será aplicada no registro jurídico escolhido.

A lei de Bradford foi criada por Samuel Clement Bradford (1878-1948), quando observou que alguns periódicos produziam muitos artigos, outros produziam alguns artigos e muitos, poucos artigos. Guedes (2012, p. 81) explica que a lei “permite estimar o grau de relevância de periódicos em uma área do conhecimento, que os periódicos que produzem maior número de artigos sobre o assunto formam um núcleo de periódicos, supostamente de maior qualidade ou relevância para aquela área”. Já a Lei de Lotka é relacionada à produtividade de autores, considerando o número de artigos publicados. A lei parte da afirmativa que poucos autores produzem muito e muitos autores produzem pouco.

E por fim a lei de Zipf, que permite estimar as frequências de ocorrência das palavras, buscando palavras com maior relevância dentro de determinado texto. As fórmulas criadas por Zipf deixaram lacunas, que foram trabalhadas por Goffman, enriquecendo a lei. Goffman criou uma fórmula que calcula o ponto de transição, ou seja, região que estariam as palavras de maior carga semântica de um texto em análise. A fórmula é demonstrada abaixo:

$$(8) \ n = \frac{-1 + \sqrt{1 + 8I_1}}{2}$$

Goffmann denominou n como o ponto de transição (T) e LI como a quantidade de vezes que palavras de única ocorrência aparecem.

Mas como este estudo de Goffmann, também foram apresentados outros no mesmo sentido, como o de Lapa e Corrêa (2011) que adaptou a fórmula de calculo do Ponto T de Goffman para uma fórmula que é capaz de calcular textos curtos escritos na

língua portuguesa. Ambos os estudos serão utilizados neste trabalho, para comparação entre resultados.

2.2 REGISTRO INFORMACIONAL E METADADOS

Registros de informação ou informacional são todo e qualquer tipo de informação que esteja escrita ou registrada em algum lugar. Segundo Mota (2009, p. 39):

“A informação é a chave do processo para um melhor entendimento do mundo por meio da criação de novos saberes. Nesta perspectiva o registro, representação e recuperação dos fatos/eventos (em bases documentais eletrônicas ou não) se tornam imprescindíveis à geração e gestão do conhecimento em todos os campos[...]” (Mota, 2009, p. 39)

Assim neste trabalho, usaremos esses registros de informação jurídica para estudar uma maneira de melhor gerir essas informações através da aplicação da infometria. São poucos os estudos que usam ou usaram até o momento registros jurídicos como objeto de estudo. Por este motivo que foi escolhido um registro jurídico em sites de tribunais para uma análise de seus metadados.

Metadados são, segundo Souza; Catarino e Santos (1997), “descrições de dados armazenados em banco de dados, ou como comumente definido dados sobre dados [...]”. Desta forma não passam de descrições de dados.

Souza, Vendrusculo e Melo (2000, p.93) ainda acrescentam que “metadado significa dado sobre o dado”. Esses metadados podem ser encontrados em qualquer tipo de registro de informação, neste caso o conjunto de metadados que trabalharemos posteriormente é composto por 20 elementos.

2.3 REGISTRO JURÍDICO E JURISPRUDÊNCIA

Torna-se importante para o presente estudo apresentar algumas definições de informação jurídica. Passos (1994, apud SANTOS E. L.; COELHO M. F. P., 2013, p.319), define informação jurídica, como:

“[...] toda unidade do conhecimento humano que tem a finalidade embasar manifestações do pensamento daqueles que lidam com a matéria jurídica, quando procuram estudar ou regulamentar situações, relações e comportamentos humanos, ou ainda quando interpretam e aplicam dispositivos legais.”

Miranda, D’Amore e Pinto (2013, P.99), ainda acrescentam que:

“a informação jurídica é originada fundamentalmente por um tripé informacional distinto: Legislação, Doutrina e Jurisprudência. Legislação é o conjunto normativo que regula a convivência social, elaborada pelo Poder Legislativo dos Municípios, Estados e União; a Doutrina é o conjunto de princípios expostos nas obras de direito, em que se firmam teorias ou se fazem interpretações sobre a ciência jurídica; e a Jurisprudência é a sábia interpretação e aplicação das leis a todos os casos concretos que se submetem a julgamento da justiça, que produz sentenças, no primeiro grau, ou acórdãos e súmulas, nos Tribunais.” (MIRANDA; D’AMORE; PINTO, 2013, p.99),

Neste trabalho utilizaremos apenas a jurisprudência como objeto de estudo, para que o campo não fique muito extenso.

Jurisprudência é o termo jurídico que designa o conjunto das decisões sobre interpretações das leis feitas pelos tribunais de uma determinada jurisdição. Para Nast (2011, apud MAXIMILIANO, 2011, p.146), a Jurisprudência possui, na atualidade, três funções muito nítidas, que se desenvolveram lentamente: uma função um tanto automática de aplicar a lei; uma função de adaptação, consistente em pôr a lei em harmonia com as ideias contemporâneas e as necessidades modernas; e uma função criadora, destinada a preencher as lacunas da lei. O acórdão se configura como um registro existente dentro da jurisprudência, ou seja, o acórdão é um tipo de decisão judiciária.

Sendo o objeto de estudo desse trabalho, o acórdão é um tipo de informação jurídica e trata-se, de uma representação, resumida, da conclusão a que se chegou, não abrangendo toda a extensão e discussão em que se pautou o julgado, mas tão-somente os principais pontos da discussão.

O acórdão como as demais decisões judiciais, deve apresentar o nome de seu relator, dos membros componentes do órgão julgador (câmara, turma, seção, órgão especial, plenário etc.), e o resultado da votação, que neste caso são alguns de seus metadados. Caso a votação não seja unânime, o voto vencido, ou seja, o entendimento divergente, mesmo que de um membro apenas órgão julgador deverá ser exposto no acórdão.

3 PROCEDIMENTOS METODOLÓGICOS

Esse capítulo apresenta os passos necessários à formalização da pesquisa, explicitando a metodologia utilizada, a identificação das fontes de dados e os procedimentos de análise.

3.1 DESCRIÇÃO

Esse estudo caracteriza-se exploratório descritivo. Descritivo porque irá apresentar as características de sites e registros jurídicos. Exploratório porque pretende avançar nas possibilidades de uso da infometria para espaços pouco explorados como o direito.

Neste sentido, propõe-se a medir o uso da informação jurídica com base em técnicas infométricas, trabalhando apenas com um recorte específico dentro do direito, que será os registros jurídicos em forma de jurisprudências (acórdãos), e que abordem apenas o tema pregão, para que o campo de estudo não fique muito amplo.

Pregão é uma modalidade de licitação, que se insere dentro do direito administrativo. Segundo Jacoby (2005, p.455):

“pregão é o procedimento administrativo por meio do qual a Administração Pública, garantindo a isonomia, seleciona fornecedor ou prestador de serviço, visando a execução de objeto comum no mercado, permitindo aos licitantes, em sessão pública, reduzir o valor da proposta por meio de lances sucessivos.” (JACOBY, 2005, p.455)

Esses registros jurídicos ou acórdãos serão extraídos de sites do TCU (tribunal de contas da União). O estudo será realizado em três etapas, sendo elas os objetivos específicos deste trabalho, que são: avaliar um site jurídico; analisar registros jurídicos como unidade informacional; e analisar os registros jurídicos a partir de técnicas infométricas.

A primeira etapa foi avaliar um site jurídico, para isso, foi feita uma observação e análise no site do Tribunal de Contas da União (TCU). Observando como é apresentada a informação, linguagem utilizada, conteúdo, recuperação da informação, entre outros aspectos.

Para isso, foram realizados testes de busca de informações, com o tema pregão, pelos diversos meios de busca que o site disponibiliza, para verificar os meios de recuperação e a apresentação de resultados.

Após essa análise, foi feita uma análise em registros jurídicos como unidades informacionais. Aqui, foi trabalhado apenas com os acórdãos. Onde foi possível particionar os acórdãos e estudar mais a fundo cada um de seus metadados.

E por fim, foram aplicados métodos infométricos, nestes registros jurídicos como: nuvens de palavras, redes de coocorrência, fórmulas e manipulação de dados em tabelas dinâmicas. Trabalhando em quatro pontos específicos: análise de palavras, análise de autoria, análise de decisões e análise de citações.

Para aplicação das técnicas o primeiro passo, foi identificar os acórdãos. Neste sentido foi feita uma busca no site do TCU, em pesquisa por assunto, conforme demonstrado anteriormente, e identificados 36 acórdãos que fossem diretamente relacionados ao tema “pregão”. Para isso, utilizados os seguintes padrões de busca: tema “licitação”; área “modalidade pregão”; subtema foi marcado “aspectos particulares do processo licitatório”.

Após a escolha dos acordãos, os dados foram tabulados, em uma tabela Excel, com os seguintes campos: número do acórdão, colegiado, data, relator, ementa (conteúdo inteiro), acordo (também com texto completo), e voto. Os votos foram tabulados em números 0 ou 1, onde zero significa uma decisão improcedente e 1 uma decisão procedente.

Após a tabulação dos dados os campos referentes à ementa e o acordo, foram transformados em um arquivo de texto (txt) e lançados em um software (Hermetic Word Frequency Counter Advanced), que é responsável por contar a frequência de palavras. A partir da lista de palavras foram aplicadas as fórmula de Goffman e a fórmula alternativa e elaborada uma nuvem de palavras.

A nuvem de palavras foi gerada a partir do site Wordle.net, esse site apresenta listas de palavras de uma forma criativa e visualmente bonita, de forma as palavras são diferenciadas de acordo com a frequência que aparecem e isso altera o tamanho de sua visualização

O segundo passo foi realizar uma pesquisa livre. Essa pesquisa foi realizada com os seguintes critérios de busca: selecionaram-se acórdãos e decisões como tipo de documentos; pregão; modalidade de licitação; atestado; e após o resultado filtrou-se por apenas acórdãos;

A partir desta pesquisa obteve-se 212 resultados, todos com tema pregão. A partir daí, os dados foram tabulados novamente (com os mesmos critérios anteriores) e manipulados através de planilhas dinâmicas. Foram trabalhadas questões de autorias com decisões e também de decisões ao longo do tempo.

O próximo passo foi salvar apenas as ementas dos acórdãos extraídos da pesquisa livre, desta vez, e salva-lo como arquivo de texto (txt) e novamente rodar no software. Os resultados obtidos de palavras também foram transformados em nuvens de palavras e redes de coocorrência.

Para trabalhar os acórdãos isoladamente, foram baixados os acórdãos: AC_817_2005 – Primeira Câmara, AC_1329_2006 – Plenário e AC_2482_2007 – Plenário. Foram extraídas suas ementas e transformadas em um arquivo txt, para rodar no software. Após essa fase os dados foram manipulados a fim de encontrar uma forma melhor forma de apresentar a informação.

Referente à análise de citação, foi selecionada uma súmula qualquer no site do TCU para fazer a aplicação do último tópico. Essa súmula foi obtida através da pesquisa livre, sem utilizar a nova solução, que foi detalhada na seção de verificação dos sites. Foi selecionada apenas a opção súmula para esta pesquisa, e no campo termo de pesquisa foi digitado “pregão”. A partir daí obteve-se dois resultados, um deles foi à súmula utilizada neste trabalho.

Os dados então foram tabulados com os seguintes campos: precedentes (que traz o registro jurídico precedente), precedentes sem o número, tipo (tipo do documento), código do tipo, número, ano, documento citado. O resultado final foi de 156 registros.

Com os registros tabulados, foram manipulados os dados em tabelas dinâmicas novamente, a fim de relacionar os documentos citados por ano, os tipos documentais mais citados e verificar as referências mais citadas. A partir desta tabela foi possível criar redes de coocorrência no software UCINET.

Por fim, como os resultados referente a análise de palavras estavam muito amplos (com assuntos diversos) foi realizada uma pesquisa mais refinada, na pesquisa por assunto, através dos seguintes parâmetros: área “licitação”, tema “modalidade pregão”, subtema “bens e serviços em comum”, título “caracterização de bens e serviços comuns”.

Foram obtidos 27 resultados, através desses resultados, foram extraídas suas ementas e transformadas em arquivos txt. Assim foi gerada a contagem de palavras para a ementa por arquivo. Através desta contagem foi possível gerar uma nuvem de palavras e uma matriz 2 mode e converte-la em 1 mode, para gerar a rede de coocorrência.

Depois de pronta a rede, foi feitos alguns filtros, com base na centralidade de grau degree (quantidade de relações) e quantidade de ocorrência da palavra. Considerou-se o degree entre 100 e 250. Após esse procedimento foi eliminada as palavras inúteis, ou seja, aquelas palavras que não demostram de forma alguma o possível assunto dos acórdãos estudados e elaborada uma nuvem de palavras.

Em todo o estudo foi utilizada a lei de Zipf e o ponto de transição de Goffmann, como já foi abordado em capítulos anteriores.

4 ANÁLISE DOS DADOS OBTIDOS

Esta seção aborda a análise dos dados, sendo divididos em quatro sub seções para melhor entendimento.

4.1 AVALIAÇÃO DE SITES JURÍDICOS

Neste capítulo será feita uma avaliação no site do Tribunal de Contas da União, observando pontos como: apresentação da informação, conteúdo, busca da informação, recuperação da informação.

4.1.1 TRIBUNAL DE CONTAS DA UNIÃO

Para iniciar a análise foram realizados testes no campo de acesso rápido do site, procurou-se pela letra J de “jurisprudência”, e foi possível acessar a página de jurisprudência. Outro método encontrado também na página inicial foi direto no link de jurisprudência, conforme é demonstrado na imagem 1:

Imagem 1 - – Pagina inicial do site do Tribunal de Contas da União:



Fonte: Site do Tribunal de Contas da União (2014).

Logo após clicar em jurisprudências, o usuário é encaminhado para outra página onde são exibidos alguns tipos de busca, são eles: pesquisa por assunto, pesquisa livre, pesquisa por campos, por dispositivos legais, navegação por assunto e novo sistema de busca. Abaixo segue imagem 2, para ilustrar:

Imagem 2 – Página de Jurisprudência



Fonte: Site do Tribunal de Contas da União (2014).

O próximo passo foi analisar cada uma das funções acima, para descobrir como funcionam. Iniciou-se pela pesquisa por assunto.

Pesquisa por assunto: Depois do acesso a página há uma breve explicação na parte superior em letras miúdas, onde pode-se observar que a busca é feita por súmulas.

Mais abaixo há um desenho de uma lupa representando um campo onde é possível escolher qual das colunas o usuário deseja pesquisar (dentre as opções existe: área, tema, subtema, título e entendimento). Ao lado deste símbolo existe um campo de busca por palavra chave e um botão de pesquisa.

Também existe um campo para que o usuário escolha a quantidade de registros que serão apresentados em uma página, esse campo é denominado “linhas”, há opções de exibir de 1 até “todos”. Ao lado deste existe um campo denominado “ações”, clicando nele há diversas opções, são elas: selecionar colunas, filtro, linhas por página, formato, flashback, ajuda e efetuar download. Pode-se observar melhor a partir da imagem 3:

Imagem 3 – Busca pesquisa por assunto:

← → ↻ <https://contas.tcu.gov.br/pls/apex/f?p=175:6>

TCU TRIBUNAL DE CONTAS DA UNIÃO

Jurisprudência Sistematizada Última atualização: 21/08/2014

[NAVEGAÇÃO EM ÁRVORE](#) | [PESQUISA](#) | [ASSUNTOS, RESENHAS E SÚMULAS](#) | [RESPOSTAS A CONSULTAS](#) | [INCIDENTES DE UNIFORMIZAÇÃO](#) | [SAIBA MAIS](#)

[Assuntos, Resenhas e Súmulas](#) | [Súmulas](#) | [Súmulas Revogadas](#)

FOLDER PARA CONSULTA RÁPIDA: Jurisprudência Sistematizada (SISJUR): imprima em uma única folha e mantenha ao alcance da mão.

Pesquisa por Dispositivo Legal (lei, artigo, parágrafo, etc.): CLIQUE AQUI!!! Ou clique na aba PESQUISA.

Digite uma **ÚNICA PALAVRA**(inclusive acentuação) no campo de pesquisa em branco abaixo; em seguida, clique no botão [Pesquisar].

Para visualizar resenhas (resumos), súmulas e precedentes, clique na **LENTE DE AUMENTO** posicionada à esquerda em cada linha do relatório.

> Uma nova palavra digitada no campo de pesquisa terá o efeito de REFINAR A PESQUISA ANTERIOR.
 >> Para "DESABILITAR/HABILITAR" um filtro, desmarque ou marque o pequeno quadrado localizado à direita em cada linha.
 >> Para "LIMPAR" (excluir) um filtro, clique no pequeno funil verde com "X" vermelho.

 Linhas 15

1 - 15 de 1526 

	Area	Tema	Subtema	Título	
	RESPONSABILIDADE	Pessoa física vinculada à administração	Pregoeiro e equipe de apoio	Publicidade no pregão eletrônico	-
	RESPONSABILIDADE	Pessoa física vinculada à administração	Pregoeiro e equipe de apoio	Sujeição à jurisdição do TCU	-

Fonte: Site do Tribunal de Contas da União (2014).

Foram realizados testes nestes campos, através de uma busca por acórdãos de pregões. Assim, foi escolhido o campo representado pela lupa, e selecionado a opção “pesquisar por todas as colunas”. No campo seguinte foi digitado “pregões”. O campo ações possui diversas sub menus para escolha, a primeira delas é a opção de selecionar a coluna de busca (mesma funcionalidade do campo “lupa”), por este motivo não foi selecionada.

A segunda opção dentro desse sub menu foi “filtro”, onde é possível filtrar por linha ou coluna (repetida pela terceira vez), e também por Operadores booleanos, que são palavras que têm o objetivo de definir para o sistema de busca como deve ser feita a combinação entre os termos ou expressões de uma pesquisa. Algumas das opções de operadores apresentadas são: em, igual, diferente, contido, não contido, como, nulo, não é nulo, entre outras. Aqui também há a possibilidade de escrever a expressão para os operadores. No entanto não foi marcado nada.

A terceira opção é “linhas por página” que também aparece pela segunda vez. Como já havia sido marcada, foi seguido em frente.

A quarta opção é o “formato”, dentro dessa opção existem varias divisões, a primeira delas é a classificação, onde há opções de classificar por colunas, direção (crescente/decrescente) ou classificação nula (nulos em primeiro/nulos no fim). Sendo assim, foi escolhido: classificar todas em ordem crescente, apresentando primeiro o tema, seguido de título, entendimento, subtema, área, e foi escolhido “nulos sempre no final”.

A segunda divisão é quebra de controle, opção que permite quebrar as colunas escolhida anteriormente, a fim de não aparecer alguma delas. Não foi marcada. A terceira divisão é “destacar”, permite destacar a palavra que o usuário deseja, a fim de encontrar mais rápido o que se procura. Essa opção foi marcada, para mostrar cor vermelha na célula pregão.

A quarta divisão é “flashback” que permite exibir os dados conforme existiam em um momento anterior, sendo que o usuário pode escolher a quantidade de minutos. Não foi aplicado. A quinta subdivisão é ajuda, possui o manual de cada função. E por fim a sexta função é “download” permite salvar o relatório de pesquisa, apenas com as informações de busca, e não de resultado. Oferece apenas o formato Excel.

Assim depois de marcadas as opções, foi realizada a pesquisa, e não foi obtido nenhum resultado. Então foi realizado o teste novamente sem filtros desta vez, mas novamente não foi obtido nenhum resultado.

Assim, foi testado a opção de selecionar por colunas, conforme demonstrado abaixo, na imagem 4:

Imagem 4 – Busca por Colunas

> Uma nova palavra digitada no campo de pesquisa terá o efeito de REFINAR A PESQUISA ANTERIOR.
 >> Para "DESABILITAR/HABILITAR" um filtro, desmarque ou marque o pequeno quadrado localizado à direita em cada linha.
 >> Para "LIMPAR" (excluir) um filtro, clique no pequeno funil verde com "X" vermelho.

1 - 1526 de 1526

Area	Tema	Subtema	Titulo	Entendimento
CONTRATAÇÃO DIRETA	Concurso público/processo seletivo	Admissão anterior à data da publicação da homologação do concurso/processo seletivo	-	
CONTRATO	Concurso público/processo seletivo	Admissão de servidor público posterior à data de validade do concurso	-	
CONVÊNIO E CONGÊNERE	Admissão	Admissão temporária de professor substituto	A admissão temporária de professor substituto far-se-á exclusiva de carreira, decorrente de exoneração ou demissão, falecimento ou afastamento para capacitação e afastamento ou licença de concessão em hipóteses diferentes destas.	
FINANÇAS PÚBLICAS	Admissão	Admissão temporária de professor substituto	-	
LICITAÇÃO	Concurso público/processo seletivo	Agentes da área de saúde	É exigido processo seletivo público para admissão de Agente Cc	
OBRA E SERVIÇO DE ENGENHARIA	Admissão	Amparo em decisão judicial de	-	
OUTRAS ÁREAS				
PESSOAL				
PLANEJAMENTO DA CONTRATAÇÃO				
PRINCÍPIO				
PROCESSUAL				
REGULAMENTO				

Fonte: Site do Tribunal de Contas da União (2014).

O filtro foi feito a partir das próprias colunas, em uma nova tentativa de obter resultados. Na primeira coluna foi selecionado a opção “licitação” como área, no tema “modalidade pregão”, e em subtema foi marcado “aspectos particulares do processo licitatório”. Como o campo já estava bem restrito parou-se de filtrar. Então, obteve-se 6 resultados de títulos, conforme imagem 5:

Imagem 5 – Resultados de linhas: pesquisa por assunto

1 - 6 de 6

Area	Tema	Subtema	Titulo	Entendimento
LICITAÇÃO	Modalidade pregão	Aspectos particulares do processo licitatório	Aceitabilidade dos recursos pelo pregoeiro	-
LICITAÇÃO	Modalidade pregão	Aspectos particulares do processo licitatório	Divulgação do orçamento estimado	-
LICITAÇÃO	Modalidade pregão	Aspectos particulares do processo licitatório	Negociação de preços pelo pregoeiro	-
LICITAÇÃO	Modalidade pregão	Aspectos particulares do processo licitatório	Particularidades da proposta em pregão	-
LICITAÇÃO	Modalidade pregão	Aspectos particulares do processo licitatório	Prazos relacionados ao pregão	-
LICITAÇÃO	Modalidade pregão	Aspectos particulares do processo licitatório	Solicitação de amostras em pregão	-

Fonte: Site do Tribunal de Contas da União (2014).

Desta forma, foi feita uma investigação dentro de cada linha (ou título), para verificar se o título realmente relacionava-se com o assunto do documento, começando pela primeira linha.

No apêndice B é tabulado os dados de resultados obtidos, com a quantidade de vezes que o termo de busca aparece dentro de cada acórdão obtido, separado por seus campos. Por exemplo: No acórdão 1542/2014_Plenário, é possível visualizar que a palavra “pregão” apareceu 1 vez no campo sumário, nenhuma vez no campo Relatório, e 1 vez nos campos Voto e Acórdão. Com os dados tabulados é possível perceber se a busca foi feita de forma inteligente pelo sistema. Todos os resultados que aparecem em linhas de assuntos foram simulados. Aqui percebe-se que não existe nenhuma forma para salvar o documento, apenas opção para imprimir.

Para finalizar, foi realizado um último teste, desta vez usando apenas “pregão” e não “pregões”, foi selecionado a opção “em todas as colunas” no campo lupa e clicou-se em pesquisar. Desta vez, obteve-se 23 resultados, sendo 19 pertencentes à área de licitações, 2 a área de tecnologia da informação, 1 a responsabilidade e 1 a planejamento da contratação. Assim observa-se que a busca estava sendo feita pela palavra incorreta.

Pesquisa livre: Nesta página a forma visualização é diferente (horizontal), assim como os campos. São eles: termo de pesquisa (para este existe um campo específico para selecionar em qual tipo de documento se deseja pesquisar); expressão de pesquisa; operador de pesquisa; número de documentos por página e base de dados.

O termo de pesquisa consiste em argumentos de pesquisa, ou seja, os grupamentos de caracteres utilizados como referência para a recuperação de documentos (p. ex.: “contas”, “petróleo”, “AC-0001-01/98-P”, “21/02/1997”, “1.578.009/2000-23”), expressão de pesquisa são palavras, podem ser usadas juntas com os operadores booleanos.

Número de documentos por página é a quantidade de documentos que podem ser exibidas por página, neste caso as opções são de 10 até 100, sendo que há um intervalo de 10 em cada opção.

E por fim em base de dados, são os tipos de documentos existentes na jurisprudência, são eles: acórdãos, decisões, relações, atas, súmulas, normas, processos,

contas do governo, atos de admissão e concessão. Há opção de marcar uma base, mais de uma ou todas.

No entanto, logo que foi feito o acesso a página, foi possível perceber que o sistema esta em processo de desativação, através de uma mensagem e um link para uma nova solução de busca na parte superior da página. Conforme ilustra a imagem 6:

Imagem 6: Mensagem de Desativação de Solução de Busca

The screenshot shows the TCU search system interface. At the top, there is a header with "Pesquisa Livre" and "Coletânea". Below this, a red message states: "Sistema em processo de desativação." (System in process of deactivation). Below the message, it says: "Caso tenha alguma dúvida, crítica ou sugestão, escreva para pesquisa@tcu.gov.br." and "Acesse a nova solução de busca do TCU." (Access the new search solution of the TCU).

The search interface includes a search bar with "Termo de pesquisa" and "Expressão de pesquisa" fields, a dropdown menu for "No campo...", and a dropdown menu for "Operador de pesquisa". There are buttons for "Pesquisar" and "Limpar". Below the search bar, it says "Documentos por página: 20".

On the right side, there is a section titled "Bases de dados" with a list of checkboxes for different data sources:

- ☒ Acórdãos
- ☐ Decisões
- ☐ Relações
- ☐ Atas
- ☐ Súmulas
- ☐ Todas as bases
- ☐ Normas
- ☐ Processos
- ☐ Contas do Governo
- ☐ Atos de Admissão e Concessão

At the bottom, there is a table with the following data:

Opções	#	Refinando	Documentos	Termos	Base	Expressão	?
<input checked="" type="checkbox"/>	2		109	1777	Acórdãos	pregão E sistema s	
<input checked="" type="checkbox"/>	1		3917	65050	Acórdãos	pregão	

Below the table, there is a footer with the following text:

>> Em caso de dúvidas, críticas e sugestões, favor entrar em contato: [Jurisprudência](#)
 >> Requisição atendida em 0.502 segundo(s).

Fonte: Site do Tribunal de Contas da União (2014).

Por este motivo, será usado para fins de estudo a nova solução, que já esta disponível. Esta solução nova de busca é chamada de consulta textual, com poucos campos de busca. São eles: um campo de múltipla escolha, para escolher o tipo de documento, um campo de pesquisa livre (texto) e mais abaixo, caso o usuário tenha interesse existe um campo de busca avançada.

Foi realizado um teste de pesquisa pela palavra “pregão” e selecionado a opção acórdãos, na busca simples. No entanto obteve-se 3.883 resultados.

A solução mostra junto aos resultados opções de filtro por relatores, e colegiado. Então foi realizada uma nova busca através da pesquisa avançada. A pesquisa foi feita conforme parâmetros mostrado na imagem 7:

Imagem 7- Parâmetros de Pesquisa

Pesquisa avançada em: Acórdãos e decisões Voltar

Localizar documentos com os seguintes argumentos

todas estas palavras:

esta expressão ou frase exata:

qualquer uma destas palavras:

nenhuma destas palavras:

números que variam de: a

Pesquisar Limpar

Fonte: Site do Tribunal de Contas da União (2014).

Obteve-se a partir da pesquisa 108 resultados. Sendo que desses pode-se filtrar ainda por relevância, data, tipo e título.

Ainda na pesquisa avançada, existe mais uma forma de busca, essa forma consiste em montar seus argumentos de pesquisa, através de campos: campo de pesquisa, termo de pesquisa e operadores booleanos. Conforme segue na Imagem 8:

Imagem 8 – Pesquisa avançada por argumentos

OU

Montar seu argumento de pesquisa

> >

Argumento de pesquisa gerado

Fonte: Site do Tribunal de Contas da União (2014).

Aqui a pesquisa foi realizada pelos mesmos termos da pesquisa anterior, utilizando os operadores booleanos E, e obteve-se 978 resultados/documentos.

Na primeira pesquisa ainda foi reduzido o volume de documentos através de um filtro apenas por acórdãos e não acórdãos relacionados, e destes somente acórdãos da primeira câmara, assim obteve-se 18 resultados. Estes foram tratados e tabelados, da mesma forma que os acórdãos por assunto. Neste caso foram 3 tabelas, um para cada termo de pesquisa, conforme Apêndice C.

Pesquisa por dispositivo legal: Este é um tipo de pesquisa mais restritivo de certo modo, pois o usuário precisa saber o que se deseja pesquisar em termos de dispositivos legais, como excertos de acórdãos ou decisões por referência legal (critério de auditoria, ato normativo, etc). Um leigo, por exemplo, não conseguiria pesquisar através deste método.

Dentre os campos de pesquisa estão: órgão emissor, norma, número da norma, ano, artigo, parágrafo, inciso, alínea. Em órgão emissor, há a possibilidade de escolher uma entidade, assim como no campo norma, existem várias opções dentre elas estão: acordo, leis, decretos, mandado de segurança, portaria, entre outras. Em número da norma, ano e artigo há um campo para descrição dos números. E por fim os campos: parágrafo, inciso e alínea também são de múltipla escolha.

O primeiro teste foi um acordo de número 686/2011, do TCU, que aborda as diversas condutas adotadas pelos responsáveis pelas licitações. Neste sentido, no primeiro campo foi selecionado “tribunal de contas da união”, no segundo campo selecionou-se “acordo”, e nos dois seguintes “686” e “2011”. Os quatro últimos não nos interessam, portanto não serão incluídos. Desta forma não foi obtido nenhum resultado.

Para o segundo teste, foi usado um acórdão que foi obtido como resultado na pesquisa anterior (pesquisa livre), trata-se do acordo 288/2007 - Plenário. Pode-se assim, afirmar que esse acordo encontra-se na base do TCU. Então usando os mesmos parâmetros da primeira tentativa, apenas mudando o numero e ano do acordo, fez-se a pesquisa e novamente não houve dados recuperados.

A terceira tentativa, foi buscar a lei de licitações 8.666/1993, para verificar se realmente o problema era no sistema de busca ou nos acórdãos buscados que não

estavam cadastrados. Foram preenchidos os campos com: presidência da república, decreto lei, numero e ano da lei. Não foi obtido nenhum resultado. Então, novamente foi feita a busca por congresso nacional, lei ordinária, numero e ano da lei, e artigo 37. Neste, obteve-se quatro resultados de acórdãos que citam o artigo 37 da lei. No entanto, não existem métodos para salvar o documento.

Pesquisa por campos: Este método de busca foi acessado no dia 17/09/2014, e assim como o de pesquisa livre encontra-se em processo de desativação, por isso não será estudado. Segue imagem 9, que ilustrativa a mensagem:

Imagem 9 – Mensagem de desativação pesquisa por campos

Pesquisa em Formulário

Sistema em processo de desativação.

Caso tenha alguma dúvida, crítica ou sugestão, escreva para pesquisa@tcu.gov.br.

Acesse a nova solução de busca do TCU.

Tipo:

Nº do Documento:

Colegiado:

Nº do Processo:

Relator:

Pesquisa Livre:

Acórdãos

?

▼

▼

▼

Pesquisar

Limpar

Ano:

Ano:

Documentos por Página:

20

 ▼

Fonte: Site do Tribunal de Contas da União (2014).

Navegação por assunto: Este método de busca funciona em forma de árvore, ou seja, quando se clica em um assunto ele mostra várias opções para escolha. Em sua página inicial na parte superior da página, aparece um campo para escolher a visualização da árvore, onde existem três opções, são: Obras e serviços de engenharia, tecnologia da informação ou todos. Mais abaixo pede para escolher a área de interesse da pesquisa. Conforme imagem 10:

Imagem 10 – Pesquisa em Árvore

NAVEGAÇÃO EM ÁRVORE	PESQUISA	ASSUNTOS, RESENHAS E SUMULAS	RESPOSTAS A CONSULTAS	INCIDENTES DE UNIFORMIZAÇÃO	SAIBA MAIS
Ordem de afinidade Ordem alfabética					

Todas as áreas estão presentes na árvore

Novo!!! ===== Incidentes de Uniformização de Jurisprudência: CLIQUE AQUI!!!! Ou clique na aba IUJ. ===== **Novo!!!**
Novo!!! ===== Respostas a Consultas: CLIQUE AQUI!!!! Ou clique na aba RESPOSTAS A CONSULTAS. ===== **Novo!!!**
Novo!!! ===== PESQUISA RÁPIDA: CLIQUE AQUI!! Ou clique na aba ASSUNTOS, RESENHAS E SUMULAS. ===== **Novo!!!**

Recolher Lista

Opções Visualização Árvore

Todos

Obra e Serviços de Engenharia

Tecnologia da Informação

Todos

Escolha a Área de Interesse

PLANEJAMENTO DA CONTRATAÇÃO

LICITAÇÃO

CONTRATAÇÃO DIRETA

CONTRATO

LICITAÇÃO INTERNACIONAL

CONVÊNIO E CONGÊNERE

OBRA E SERVIÇO DE ENGENHARIA

TECNOLOGIA DA INFORMAÇÃO

PUBLICIDADE E PROPAGANDA

Fonte: Site do Tribunal de Contas da União (2014).

Para fins de teste, foi escolhida a área de interesse “licitação”. Após a escolha é requerida mais uma opção de escolha que se trata do assunto que se deseja, neste foi selecionada a opção “modalidade pregão”. Logo após é requerida a terceira opção de escolha referente a o que se deseja buscar neste assunto, muito semelhante ao primeiro sistema de busca (pesquisa por assunto). Foi selecionada aqui a primeira opção, que é “aspectos particulares do processo licitatório”, que foi a opção escolhida no método de busca por assunto. Isso para verificar se os resultados serão os mesmo. Abaixo segue imagens 11 e 12, com a ilustração:

Imagem 11 – Visualização do sistema de busca em árvore: primeira divisão

UVV!!! ----- PESQUISA RÁPIDA: CLIQUE AQUI!!! OU CLIQUE NA ABA ASSUNTOS, RESOLUÇÕES E SUMULAS. ----- NOVU!!!

Recolher Lista

Opções Visualização Árvore Todos ▼

- ↳ Dimensionamento
- ↳ Modalidade convite
- ↳ Modalidade tomada de preços
- ↳ Modalidade concorrência
- ↳ Modalidade pregão
- ↳ Registro de preços
- ↳ Tipo de licitação
- ↳ Formalização de processo licitatório
- ↳ Publicidade dos atos
- ↳ Ato convocatório
- ↳ Aprovação de ato convocatório e anexos
- ↳ Audiência pública
- ↳ Participação e competitividade
- ↳ Prazos
- ↳ Proposta e julgamento
- ↳ Amostra ou protótipo

Fonte: Site do Tribunal de Contas da União (2014).

Imagem 12 - Visualização do sistema de busca em árvore: segunda divisão

Recolher Lista

Opções Visualização Árvore Todos ▼

- ↳ Aspectos particulares do processo licitatório
- ↳ Bens e serviços comuns
- ↳ Eletrônico
- ↳ Obra e serviço de engenharia
- ↳ Presencial
- ↳ Uso obrigatório
- ↳ Modalidade Pregão
- ↳ Registro de preços
- ↳ Tipo de licitação
- ↳ Formalização de processo licitatório
- ↳ Publicidade dos atos
- ↳ Ato convocatório
- ↳ Aprovação de ato convocatório e anexos
- ↳ Audiência pública
- ↳ Participação e competitividade
- ↳ Prazos

Fonte: Site do Tribunal de Contas da União (2014).

Neste são apresentados os mesmos resultados já obtidos e tabulados, com exceção de dois assuntos (imagem 13), que são “licitação/ato convocatório/anexos – demonstrativo do valor estimado” e “outros procedentes não consolidados”. Sendo que o primeiro apresenta 57 resultados enquanto o segundo apresenta 89.

Imagem 13 – Apresentação dos assuntos de resultados obtidos

Novo!!! ==== Incidentes de Uniformização de Jurisprudência: CLIQUE AQUI!!!! Ou clique na aba IJ, ==== Novo!!!
 Novo!!! ==== Respostas a Consultas: CLIQUE AQUI!!!! Ou clique na aba RESPOSTAS A CONSULTAS. ==== Novo!!!
 Novo!!! ==== PESQUISA RÁPIDA: CLIQUE AQUI!! Ou clique na aba ASSUNTOS, RESENHAS E SÚMULAS. ==== Novo!!!

Recolher Lista

Opções Visualização Árvore Todos ▼

- Assunto não consolidado: PARTICULARIDADES DA PROPOSTA EM PREGÃO
- Assunto não consolidado: SOLICITAÇÃO DE AMOSTRAS EM PREGÃO
- Assunto não consolidado: PRAZOS RELACIONADOS AO PREGÃO
- Assunto não consolidado: ACEITABILIDADE DOS RECURSOS PELO PREGOEIRO
- Assunto não consolidado: DIVULGAÇÃO DO ORÇAMENTO ESTIMADO
- Assunto não consolidado: NEGOCIAÇÃO DE PREÇOS PELO PREGOEIRO
- Assunto não consolidado: LICITAÇÃO / ATO CONVOCATÓRIO / ANEXOS - DEMONSTRATIVO DO ORÇAMENTO ESTIMADO
- OUTROS PRECEDENTES NÃO CONSOLIDADOS
- TODOS OS PRECEDENTES DO SUBTEMA

Fonte: Site do Tribunal de Contas da União (2014).

Neste sistema também não há formas de salvar os documentos gerados.

Assim pode-se concluir com essa análise que o site apresenta vários sistemas de busca, no entanto com certa dificuldade e complexidade de uso, existem palavras ou expressões específicas que exigem conhecimento sobre os métodos.

Ainda pode-se observar que há uma limitação no que diz respeito a recuperação dessa informação. Levando em consideração que há apenas alguns formatos disponíveis para salvar essa informação, e em muitas formas de busca não se consegue salvar a informação em lotes (mais de um arquivo ao mesmo tempo) e ainda algumas destas formas nem possuem opções para salvar os dados.

Devido ao site não disponibilizar uma forma de recuperação em lote das informações, o tratamento fica mais complicado, sendo que para tal é necessário baixar um por um dos documentos existentes. O que caracteriza uma lacuna muito grande no site.

A informação que é recuperada também nem sempre é o que se busca, devido a assuntos muito diferentes estarem acoplados, o que faz com que se perca muito tempo a procura da informação correta.

Finalmente é possível afirmar que a informação disponível nos sites poderiam ser melhor tratadas a fim de ajudar o usuário em sua busca rotineira.

4.2 ACÓRDÃO

A partir de agora será estudado apenas o acórdão e as partes que o compõe (chamados metadados), a fim de mostrar que esse documento pode ser estudado enquanto registro de uma informação.

O acórdão foi escolhido como documento a ser estudado devido a sua importância dentro da área jurídica, caracterizado como um tipo de decisão, e também por apresentar vários campos (metadados) que oferecem meios de ser trabalhados.

Na primeira parte da tabela abaixo se tem o nome do campo ou metadado e em seguida a explicação do campo, também é ilustrado a partir de um acórdão qualquer.

Metadado	Explicação
Número interno do documento	Número interno identificador do documento

Número Interno do Documento:
AC-7498-44/12-1

Metadado	Explicação
Colegiado	Colegiado a que pertence, pode ser: Plenário, Primeira Câmara ou Segunda Câmara.

Colegiado:
Primeira Câmara

Metadado	Explicação
Relator	Relator responsável

Relator:
JOSÉ MÚCIO MONTEIRO

Metadado	Explicação
Processo	Número do processo

Processo:
021.466/2008-1

Metadado	Explicação
Sumário	Breve relato dos fatos

Sumário:
PRESTAÇÃO DE CONTAS RELATIVA AO EXERCÍCIO DE 2007. INDÍCIOS DE IRREGULARIDADES. AUDIÊNCIA DE ALGUNS RESPONSÁVEIS. ACOLHIMENTO DAS RAZÕES DE JUSTIFICATIVA. CONTAS REGULARES COM RESSALVA DE UM RESPONSÁVEL E REGULARES DOS DEMAIS

Metadado	Explicação
Assunto	Assunto do documento

Assunto:
Prestação de contas

Metadado	Explicação
Número do acórdão	Número identificador do documento

Número do acórdão:
7498

Metadado	Explicação
Ano do Acórdão	Ano que documento foi publicado

Ano do acórdão:
2012

Metadado	Explicação
Número da data	Número de acórdão publicado daquele ano

Número da ata:
44/2012

Metadado	Explicação
Dados materiais	Quantidade de volumes e anexos do documento

Dados materiais:
Dados Materiais:(com 1 volume e 1 anexo)

Metadado	Explicação
Relatório	Relato do ocorrido

Relatório:

Trata-se de prestação de contas do Serviço de Apoio às Micro e Pequenas Empresas no Estado da Bahia (Sebrae/BA) relativa ao exercício de 2011.

2. O processo foi redistribuído para a Secex/SC e a auditora dele encarregada apresentou a instrução transcrita a seguir:

"1. Tratam os autos de prestação de contas dos responsáveis pelo Departamento Regional do Sebrae na Bahia - Sebrae/BA relativas ao exercício de 2011.

2. Após análise preliminar, esta unidade técnica realizou a audiência dos responsáveis abaixo indicados para apresentarem razões de jure e de fato.

Metadado	Explicação
Voto	Voto do ministro relator

Voto:
VOTO

Em exame a prestação de contas do Serviço de Apoio às Micro e Pequenas Empresas no Estado da Bahia (Sebrae/BA), relativa ao exercício de 2007.

2. Ao contrário da unidade técnica e do Ministério Público junto ao TCU, dos quais lamento divergir, penso que as ocorrências apuradas no processo, para motivar o julgamento pela irregularidade das contas.

Metadado	Explicação
Acórdão	Julgamento final

Acórdão:

VISTOS, relatados e discutidos estes autos de prestação de contas anual do Serviço de Apoio às Micro e Pequenas Empresas no Estado da Bahia (Sebr

ACORDAM os Ministros do Tribunal de Contas da União, reunidos em sessão da 1ª Câmara, com fundamento nos arts. 1º, inciso I; 16, incisos I e II; 17 expostas pelo Relator, em:

Metadado	Explicação
Entidade	Entidade participante ou julgada

Entidade:

Entidade: Serviço de Apoio às Micro e Pequenas Empresas no Estado da Bahia (Sebrae/BA)

Metadado	Explicação
Interessado	Interessados no processo

Interessado:

Responsáveis: Alberto Bispo do Nascimento (CPF nº 370.768.845-34); Angelo Mário de Carvalho Silva (CPF nº 002.948.535-53); Antonio Loma (CPF nº 001.118.215-68); Aristóteles Alves de Menezes Júnior (CPF nº 312.983.625-04); Carlos Fernando Amaral (CPF nº 000.644.325-72); Cl 056.158.365-04); Fernando Elias Salamoni Cassis (CPF nº 006.664.245-00); Franklin Santana Santos (CPF nº 777.000.825-53); Geraldo Afonso

Metadado	Explicação
Representante do MP	Representante do Ministério Público

Representante do MP:
Júlio Marcelo de Oliveira

Metadado	Explicação
Unidade técnica	Unidade técnica

Unidade técnica:
Secex/SC

Metadado	Explicação
Advogado	Advogado das partes

Advogado:
Sérgio Dutra Ribas (OAB/BA nº 13.903)

Metadado	Explicação
Quórum	Ministros e ministros substitutos presentes

Quórum:
13.1. Ministros presentes: Valmir Campelo (Presidente), Walton Alencar Rodrigues, José Múcio Monteiro (Relator) e Ana Arraes.
13.2. Ministros-Substitutos presentes: Augusto Sherman Cavalcanti e Weder de Oliveira

Metadado	Explicação
Data da Sessão	Data da Sessão

Data da sessão:
04/12/2012

A partir da análise anterior pode-se concluir que o acórdão pode ser trabalhado enquanto unidade informacional, já que possui vários campos textuais que permitem trabalhar seu conteúdo, assim como possui informações temporais e de autoria, o que também permite trabalhar com esses dados relacionados a decisões (campo voto e acórdão).

Cada um dos metadados acima pode ser tratado e alvo de análises infométricas, alguns deles, por exemplo, serão apresentados no próximo capítulo.

4.3 APLICAÇÃO DE TÉCNICAS INFOMETRICAS SELECIONADAS

Para melhor análise e apresentação de resultados trabalhou-se com cinco focos principais neste último tópico, conforme abaixo:

- Análise de palavras;
- Representação em nuvem;
- Análise de autoria;
- Análise da decisão
- Análise de citações - Genealogia.

No primeiro tópico (análise de palavras), trabalhou-se com acórdãos pertencentes à pesquisa por assunto e a pesquisa livre. Sendo essa dividida em quatro estudos: (i) análise da ementa de acórdãos isolados, pertencentes a pesquisa por assunto; (ii) análise de ementas, acórdãos e ementa mais o campo acórdão juntos em acórdãos de pesquisa por assunto; (iii) ementas de acórdãos obtidos pela pesquisa livre e por fim (vi) análise em ementas obtidas através de pesquisa por assunto, mais específicos do que o primeiro. Foram aplicadas duas técnicas de infometria para tratamento do conteúdo. As técnicas são:

- Contagem de palavras (na linha da Lei de Zipf);
- Nuvens de palavras.

Ainda neste tópico, foi utilizado o ponto de Transição (T) de Goffman, para encontrar as palavras que obtém maior relevância dentro de um determinado número de palavras ou texto. Para isso foi utilizado a seguinte fórmula:

$$n = \frac{-1 + \sqrt{1 + 8I_1}}{2} \quad (1)$$

Onde L^1 é o número de palavras que possui uma única ocorrência e o ponto T é igual a n. Além disso foi utilizada uma fórmula alternativa, que se trata de uma adaptação da fórmula anterior, feita por Lapa e Corrêa (2011, p. 9) sendo essa aplicada à palavras em português. A seguir é apresentado a fórmula:

$$n = \frac{-1 \pm \sqrt{1 + 36 \times L1}}{9} \quad (2)$$

Dessa forma a variável L1 corresponde ao número de palavras que têm frequência igual a um; e o número 9 (nove) uma constante atribuída à língua portuguesa.

No segundo tópico (representação em nuvem), trabalhou-se com a visualização dessa informação gerada no tópico um e buscou-se mostrar a aplicabilidade desta ferramenta na visualização da informação jurídica.

No terceiro tópico (análise de autoria), foi trabalhado com a pesquisa livre de acórdãos, onde o número de resultados foi significativamente maior. Estes resultados foram tabulados, a partir de seus metadados e manipulados, a fim de buscar algum tipo de informação sobre as decisões e os relatores. Também foram analisados aspectos referentes ao desempenho dos relatores neste tópico.

O quarto tópico (análise da decisão), também trabalhou com a pesquisa livre, e caracterizou-se como uma extensão do tópico anterior, em que foi possível verificar a incidência de deferimentos ou decisões que consideradas procedentes em relação às não procedentes. Procurou-se, através dos dados tabulados, analisar as diferentes decisões ao longo do tempo também.

E por fim, no quinto tópico (análise da citação - genealogia), foi feito um estudo em uma súmula específica que abordasse o tema pregão, e analisado as referências de acórdãos mencionados, para verificar se existe alguma relação entre seus primórdios e se é possível trabalhar melhor a informação conhecendo sua história.

4.4 ANÁLISE DA APLICAÇÃO DAS TÉCNICAS

4.4.1 ANÁLISE DE PALAVRAS

A aplicação das técnicas será analisada seguindo os tópicos no capítulo anterior. O que será feita em quatro fases. Primeiramente será feita a análise de palavras em uma ementa de três acórdãos isolados, observando sua evolução na apresentação e tratamento da informação. Ainda nesta fase esses acórdãos serão relacionados para análise. E então, o próximo passo será verificar o procedimento de análise de palavras em vários acórdãos acoplados, considerando a pesquisa por assunto e em seguida a pesquisa livre, analisando seus campos de ementas, acórdão e ambos juntos.

Na última fase é demonstrado o resultado da pesquisa por assunto de bens e serviços em comum. Essa pesquisa foi feita com assuntos mais específicos, para verificar se o resultado seria alterado.

4.4.1.1 ANÁLISE ISOLADA DE ACÓRDÃOS

Foram analisados os acórdãos AC_817_2005 – Primeira Câmara, AC_1329_2006 – Plenário e AC_2482_2007 – Plenário, como já foi mencionado. E feita a contagem de palavras, no entanto não foi possível aplicar a fórmula do ponto de transição de Goffmann devido a pequena quantidade de palavras.

Dessa forma a análise foi elaborada a partir das palavras mais citadas de cada acórdão. O primeiro a ser analisado foi o AC_1329_2006 – Plenário, onde as palavras de maior frequência foram as palavras mencionadas na tabela abaixo, com frequência 2 cada uma:

Tabela 1 – Resultado de palavras AC_1329_2006

Palavras	AC_1329-2006
Total	23
Pregão	2

serviços	2
engenharia	2
representação	2
eletrônico	2
contratação	2

Fonte: Aautora (2014).

A partir do resultado acima é possível verificar que as palavras formam um bom grupo de palavras para a indexação do acórdão, levando em consideração que o acórdão fala de uma representação feita ao tribunal de Contas, acusando uma irregularidade no processo licitatório de modalidade pregão eletrônico, para contratação de serviços de engenharia.

O próximo acórdão estudado foi o AC_817_2005 – Primeira Câmara, a lista de palavras obtidas foram da mesma forma que o acórdão anterior, sua frequência foi 3 e 2 já que somente usando a frequência 3 foi obtido apenas 2 resultados, abaixo segue tabela 2 com as palavras:

Tabela 2 – Resultado de palavras AC_817_2005

Palavras	AC 817_2005
Pregão	3
Serviços	3
Engenharia	2
Pelo	2
Que	2

Fonte: Aautora (2014).

Como o resultado apresentou entre stopwords (palavras como: do, das, pelo, que, etc) que caracterizam-se inúteis, foram retiradas para elaboração da nuvem de palavras, que será observada no próximo tópico.

A partir dos resultados é possível perceber que ainda que o conjunto de palavras sem as stopwords fique muito pequeno para uma indexação, as palavras que restaram representam muito bem o assunto do presente acórdão, que trata da contratação de serviços de engenharia através da modalidade pregão.

Ainda foi trabalhado no AC_2482_2007 – Plenário, que através dos mesmos procedimentos, foi obtida frequência 2, conforme tabela 3:

Tabela 3 – Resultado de palavras AC_2482_2007

Palavras	AC 2482_2007
Serviços	2
Engenharia	2

Eletrônico	2
------------	---

Fonte: Aautora (2014).

Neste o resultado ficou bem semelhante aos demais. Sendo que todos os acórdãos selecionados tratam-se do mesmo assunto, e suas decisões finais são decisões improcedentes.

Foi realizados um teste juntando todos os acórdãos, foi escolhida a frequência de palavras de 7 (número máximo) ao número 3. O resultado de palavras segue na tabela 4:

Tabela 4 – Resultado de palavras acórdãos juntos

Palavras	Todos
Pregão	7
Serviços	7
Engenharia	6
Eletrônico	4
Representação	4
Pelo	3
Comuns	3
Improcedência	3
Para	3
Por	3
Realização	3
Contratação	3

Fonte: Aautora (2014).

Percebe-se neste tópico que as palavras quando extraídas de um conjunto de acórdão pertencentes a um assunto específico mostram um resultado bem positivo. Onde poderiam facilmente ser usados para uma indexação. Além disso, trabalhar de uma forma isolada com acórdãos permite um resultado mais qualitativo, melhorando a análise. Observa-se ainda que é possível trabalhar com um ou mais acórdãos juntos e que vale a pena um estudo mais aprofundado de palavras.

4.4.1.2 ANÁLISE DE PALAVRAS PESQUISA POR ASSUNTO

Para esta análise foram extraídos 36 acórdãos de uma pesquisa realizada por assunto, conforme explicado na metodologia. Foram selecionados acórdãos referente ao critério de busca com área licitações, tema modalidade pregão, subtema aspectos

particulares do processo licitatório. Foram realizadas análises em três partes diferentes do acórdão: ementa, acórdão e acórdão e ementas juntos.

Em todas foram aplicadas duas fórmulas para extrair as palavras mais relevantes e consequentemente as que serão trabalhadas. As fórmulas são a do ponto de Transição (T) de Goffman e uma fórmula alternativa que segue os mesmos parâmetros mais é adaptada para palavras em português, conforme especificado no início do capítulo.

Primeiramente será trabalhado apenas com o campo ementa, no entanto esse campo será trabalhado de duas formas: a primeira sem as stopwords e sem S, ou seja, será excluídas as palavras como: de, pelo, para, das, os, etc e serão consideradas como uma só palavras que estejam no plural ou não.

Abaixo, na tabela 5, é demonstrado um conjunto de palavras encontradas para o campo ementa sem stopwords e sem S, usando a fórmula de Goffman, onde a frequência foi igual a 28 (que é a palavra em amarelo), ou seja, o que estiver em torno desta palavra, segundo a fórmula são as palavras mais relevantes:

Tabela 5 - Resultado campo ementa pesquisa por assunto 1

Frequência	Palavras
36	Representação
29	Pregão
28	Irregularidades
27	Cautelar
24	Dos
21	Parcial
21	Procedência
19	Conhecimento
18	Eletrônico
16	Determinações

Fonte: Aautora (2014).

Percebe-se que apesar do conjunto não apresentar nenhuma palavra inútil ainda fica muito vago o assunto dos acórdãos através das palavras. Isso pode acontecer devido a um número muito grande de acórdãos reunidos ou mesmo devido a temas muito diferentes. Abaixo é feita a análise no campo ementa sem stopwords e sem S com a fórmula alternativa, cuja frequência foi igual a 13, conforme tabela 6:

Tabela 6 - Resultado campo ementa pesquisa por assunto 2

Frequência	Palavras
18	Eletrônico

16	Determinações
15	Das
14	Medida
13	Arquivamento
12	Certame
11	Ciência
10	Licitação
10	Oitiva
10	Preços
10	Registro

Fonte: Aautora (2014).

Percebe-se que o resultado aqui, foi menos satisfatório, incluindo ainda palavras inúteis como oitiva. Também não é possível imaginar o assunto desses acórdãos através do conjunto acima. Abaixo será apresentado o resultado para o campo ementa com stopwords e utilizando primeiramente a fórmula do ponto de Goffman, conforme tabela 7:

Tabela 7 - Resultado campo ementa pesquisa por assunto 3

Frequência	Palavras
36	Representação
33	Para
29	Pregão
28	Irregularidades
27	Cautelar
24	Dos
21	Parcial
21	Procedência
19	Conhecimento
18	Eletrônico

Fonte: Aautora (2014).

Percebe-se que este resultado foi muito parecido ao primeiro, no entanto com acréscimo de algumas palavras inúteis como para e dos. A seguir será demonstrado o resultado utilizando a fórmula alternativa, cuja a frequência deu 14, conforme tabela 8:

Tabela 8 - Resultado campo ementa pesquisa por assunto 4

Frequência	Palavras
17	Não
16	Determinações
15	Das
15	Que

14	Medida
13	Arquivamento
12	Certame
11	Ciência
11	Com
10	Licitação

Fonte: Aautora (2014).

A partir do resultado pode-se afirmar que até o momento esse seria o pior conjunto para uma indexação, considerando o grande volume de stopwords.

Abaixo será trabalhado com o campo acórdão, esse estudo foi feito apenas utilizando as stopwords. O resultado aplicando a fórmula do T de Goffman retornou frequência 45, conforme tabela 9:

Tabela 9 - Resultado campo acórdão pesquisa por assunto1

Frequência	Palavras
50	Presente
48	Interno
48	Razões
48	Regimento
45	Contas
45	Tcu
43	Sessão
42	Bem
41	Serviços
39	Ciência

Fonte: Aautora (2014).

Observa-se que, é muito difícil através do resultado entender do que se trata os acórdãos selecionados, devido a presença de palavras inúteis, como: tcu, presente, bem. Para comparação, abaixo é apresentado o resultado com a fórmula alternativa, cuja frequência foi 21, conforme tabela 10:

Tabela 10 - Resultado campo acórdão pesquisa por assunto 2

Frequência	Palavras
23	Licitação
23	Objeto
22	Deliberação
22	Fundamentam
21	Processo
20	Administração
20	Câmara

20	Licitantes
20	Parágrafo
19	Arquivar

Fonte: Aautora (2014).

O resultado acima, apesar de não apresentar stopwords, também não foi muito satisfatório devido também a quantidade de palavras inúteis. A seguir serão apresentados agora os resultados obtidos referentes aos campos acórdão e ementa reunidos, primeiramente fazendo uso do ponto de transição de Goffman, cuja frequência para corte ficou em 47 ocorrências, conforme tabela 11:

Tabela 11 - Resultado campo ementa mais acórdão pesquisa por assunto 1

Frequência	Palavras
49	Serviços
48	Contas
48	Interno
48	Regimento
48	Tcu
43	Sessão
42	Bem
42	Mérito
41	Cautelar
40	Ser

Fonte: Aautora (2014).

Novamente o resultado obtido não foi bom, em sua maior parte com palavras que não dizem respeito a assunto algum. Utilizando a fórmula alternativa foi obtido frequência 22, o resultado é apresentado abaixo na tabela 12:

Tabela 12 - Resultado campo ementa mais acórdão pesquisa por assunto 2

Frequência	Palavras
23	Preço
23	Processo
22	Certame
22	Fundamentam
22	Licitantes
22	Procedência
21	Conhecimento
21	Proposta
21	Valores
20	Câmara

Fonte: Aautora (2014).

Aqui o resultado apresenta-se um pouco melhor, no entanto ainda é impossível saber em exato o assunto dos acórdãos estudados.

Pode-se concluir a partir dos resultados que há uma dispersão de assuntos muito grande nestes acórdãos selecionados, e que também para melhor estudo é interessante utilizar uma quantidade menor de acórdãos ou deixar a relação de acórdãos mais específicos, utilizando talvez alguma técnica complementar e/ou alternativa para refinamento dos acórdãos.

4.4.1.3 ANÁLISE DE PALAVRAS PESQUISA LIVRE

Nesta seção serão analisados apenas os resultados obtidos das ementas da pesquisa livre, que teve como critério a pesquisa pelas seguintes palavras: “licitação”, “modalidade pregão” e “atestado”, que em seu total resultou em 212 acórdãos. Primeiramente serão estudados os dados obtidos a partir do ponto de Transição de Goffman, cuja frequência de corte foi 42, conforme abaixo na tabela 13:

Tabela 13 - Resultado campo ementa pesquisa livre 1

Frequência	Palavra
48	Dou
48	Vide
47	Procedência
47	Serviços
42	Audiência
41	Certame
41	Razões
40	Não
38	Contratação
36	Justificativa

Fonte: Aautora (2014).

Percebe-se que os resultados são bem parecidos a pesquisa por assunto em questão de quantidade de palavras sem utilidade, e ainda em relação a falta de palavras que realmente possam demonstrar o assunto dos acórdãos ou grupos de assuntos. Apresentam-se abaixo as palavras obtidas através da fórmula alternativa na tabela 14:

Tabela 14 - Resultado campo ementa pesquisa livre 2

Frequência	Palavra
21	Aplicação
21	Licitações

21	Pela
21	Provimento
20	Defesa
20	Especial
20	Recomendações
19	Exigência
19	Irregularidade
18	Exercício

Fonte: A autora (2014).

As palavras acima, se parecem bastante com os resultados já obtidos nos itens anteriores. Desta forma não serviriam para uma possível indexação e também não ajudariam usuários em sua busca por alguma informação, devido a falta de palavras que indiquem o assunto.

Assim pode-se verificar que, este resultado também exige um refinamento melhor nos acórdãos para melhores resultados.

4.4.1.4 ANÁLISE DE PALAVRAS PESQUISA ESPECÍFICA POR ASSUNTO

Neste tópico foram selecionados 27 ementas de um assunto específico dentro da pesquisa por assunto, para verificar se o resultado poderia ficar melhor. Serão apresentados os resultados apenas sem stopwords. O primeiro resultado será visualizado tomando como referencia a frequência do ponto T de Gofmann, calculado como 28, conforme tabela 15:

Tabela 15 - Resultado campo ementa pesquisa específica 1

Frequência	Palavra
27	Representação
27	Arquivo
23	Pregão
21	Serviços
20	Para
16	Dos
15	Conhecimento
15	Licitação
15	Por
14	Modalidade

Fonte: A autora (2014).

Percebe-se que o resultado melhorou significativamente com uma escolha mais específica de assunto, devido a ocorrência de palavras que expressam possíveis assuntos, ao ponto que pode-se deduzir que o assunto destes acórdãos estejam ligados a serviços e que tenha alguma coisa que liga com modalidade pregão. Abaixo temos o conjunto de palavras apresentadas pela fórmula alternativa, cuja frequência deu 13, conforme tabela 16:

Tabela 16 - Resultado campo ementa pesquisa específica 2

Frequência	Palavra
15	Conhecimento
15	Licitação
15	Por
14	Modalidade
13	Arquivamento
13	Não
13	Procedência
12	Contratação
11	Que
10	Cautelar
10	Parcial

Fonte: Aautora (2014).

Aqui o resultado ficou bem próximo ao primeiro, no entanto, ainda aparecem mais palavras que não tem muita utilidade, como: parcial, cautelar. Mas ainda assim pode-se afirmar que o resultado foi bem satisfatório em relação aos dois tópicos anteriores.

4.4.2 REPRESENTAÇÃO EM NUVEM

Neste tópico será abordada a utilidade da representação em nuvem, enquanto forma de representação visual útil para representar as palavras tratadas na seção anterior ou qualquer outra parte do documento jurídico.

Assim será possível analisar os dados a partir de um método mais visual e de fácil entendimento, as nuvens de palavras. Serão utilizados os mesmos tópicos da análise de palavras.

4.4.2.1 ANÁLISE ISOLADA DE ACÓRDÃOS EM NUVEM

A partir da análise anterior foi elaborada uma nuvem de palavras mais citadas de cada acórdão (imagem 14). O primeiro a ser analisado foi o AC_1329_2006 – Plenário, onde as palavras de maior frequência foram as palavras mencionadas nos tópicos abaixo, sendo que as palavras de maior destaque foram as mais citadas (para estudo não foram utilizadas stopwords):

Imagem 14 – Resultado Acórdão 1329_2006 – Plenário

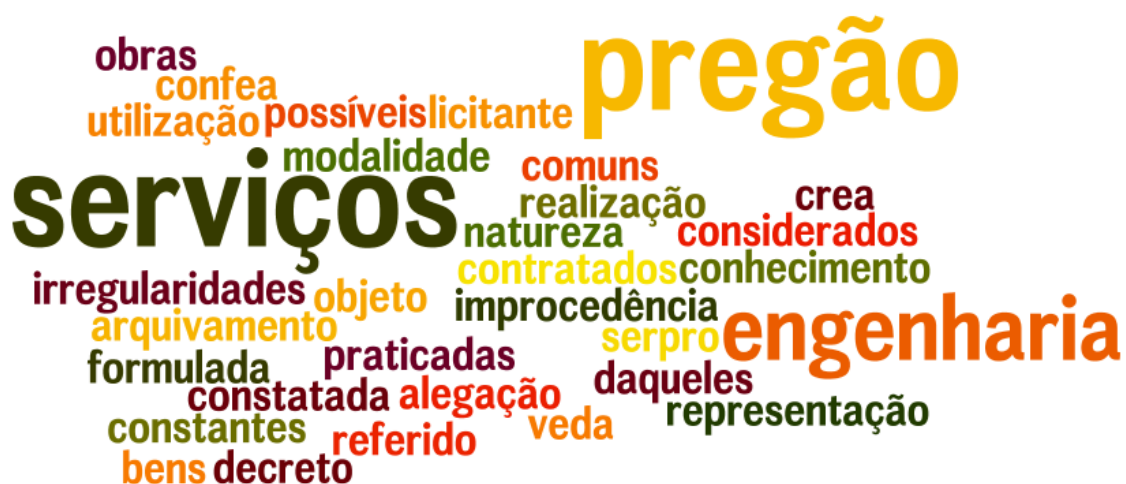


Fonte: A autora (2014).

A partir do resultado acima observa-se que as palavras representação, engenharia, pregão, eletrônico, serviços, foram as palavras que obtiveram maior frequência. Assim como fica bem claro o assunto deste acórdão a partir da imagem.

No AC_817_2005 – Primeira Câmara, a lista de palavras obtidas foram da mesma forma que o acórdão anterior, e abaixo, na imagem 15, segue a nuvem de palavras.

Imagem 15 – Resultado AC_817_2005 – Primeira Câmara



Fonte: A autora (2014).

A partir dos resultados é possível perceber que as palavras *pregão*, *engenharia* e *serviços* são as que aparecem mais vezes no texto.

Os resultados referente ao AC_2482_2007 – Plenário, seguem na imagem 16 abaixo:

Imagem 16 – Resultado AC_2482_2007 – Plenário



A autora (2014).

Percebe-se assim como no conjunto de palavras que o resultado é bem semelhante ao demais, e que todos são conjunto de dados que permitem facilmente identificar o assunto dos acórdãos.

Por fim, é apresentada abaixo a nuvem de palavras, imagem 17, desses acórdãos unidos, conforme foi feito com as palavras no tópico anterior.

Imagem 17 – Resultados dos acórdãos Reunidos



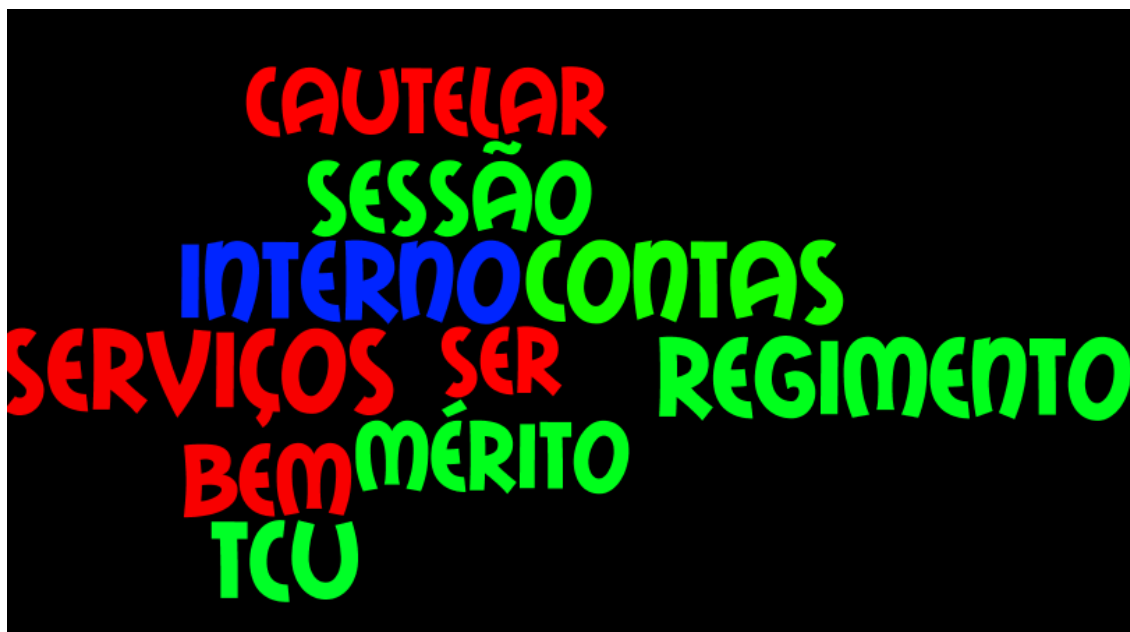
Fonte: A autora (2014).

Como nos outros o resultado aponta para uma maior relevância das palavras pregão, serviços e engenharia.

4.4.2.2 ANÁLISE EM NUVEM DE ACÓRDÃOS PESQUISA POR ASSUNTO

Abaixo é apresentado a nuvem de dados referente ao campo ementa e campo acórdão, analise a partir do Ponto de T de Goffman, conforme imagem 18:

Imagem 18 – Análise na ementa e acórdão a partir do Ponto T de Goffman



Fonte: A autora (2014).

Como foi observado anteriormente a partir do resultado percebe-se que este conjunto de palavras não seria um bom conjunto para uma indexação, levando em consideração que a maior parte delas não dizem respeito ao assunto em si, como por exemplo: plenário, regimento, TCU, ser. Abaixo foi realizado a nuvem de palavras para a fórmula alternativa. Na imagem 19 é apresentado o resultado:

Imagem 19 –Análise na ementa e acórdão a partir da fórmula alternativa



Fonte: A autora (2014).

Neste o resultado aparece um pouco melhor mais ainda assim, percebe-se que não seria uma boa forma para se encontrar alguma informação que se procura, já que o

resultado não apresenta muito bem o assunto. Abaixo também será trabalhado com o campo de acórdãos e ementas de forma isolada.

Para trabalhar com o campo de ementas, o estudo foi feito de duas formas: forma original e sem stopwords (que são palavras como: a, de, para, sem) e sem s (palavras que estão no plural). Abaixo é na imagem 20 é exibido o resultado desta segunda forma, sem stopwords e usando a fórmula de Goffmann.

Imagem 20 – Análise na ementa a partir do Ponto T de Goffman e Sem stopwords



Fonte: A autora (2014).

O resultado foi bem parecido com o que se há indexado hoje no site, possui informação sobre procedência ou não e traz o tipo de documento (representação, medida cautelar, auditoria, etc). Abaixo na imagem 21 é demonstrado a partir da fórmula alternativa:

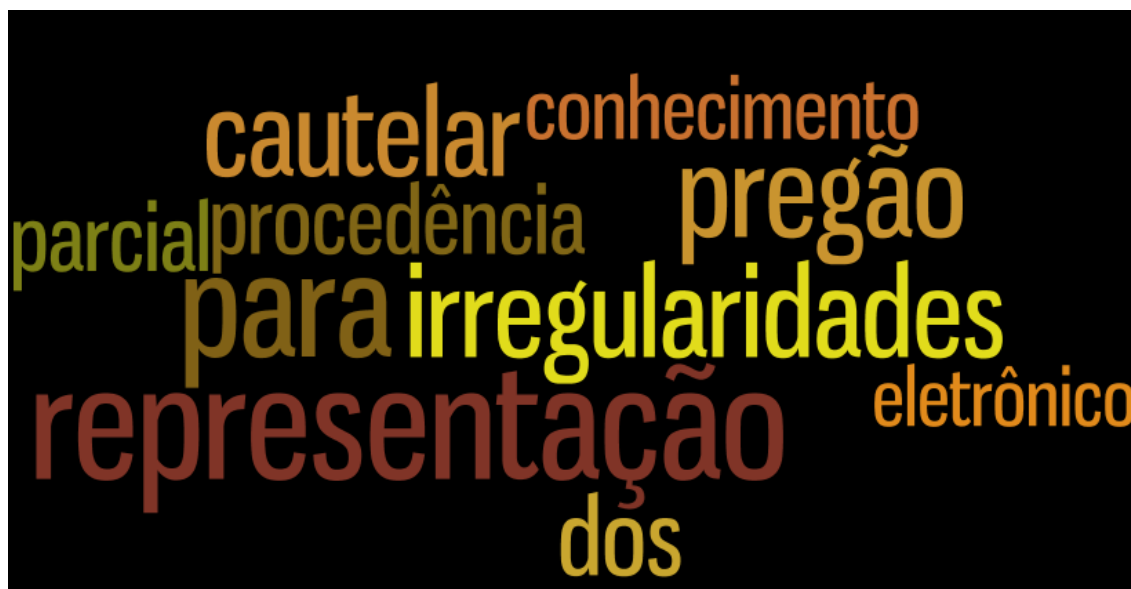
Imagem 21 – Análise na ementa a partir da fórmula alternativa e sem Stopwords



Fonte: A autora (2014).

Nesta imagem pode-se perceber que o resultado traz algumas stopwords, mas ainda assim esta bem parecido com o resultado anterior. Não traz também muitas palavras que tenham relação direta com o assunto do acórdão. Possivelmente pelo volume de acórdãos que foram acoplados como já foi comentado. Abaixo na imagem 22 será analisado o resultado referente ao campo ementa com stopwords e a partir da fórmula de Goffman:

Imagem 22 – Análise na ementa a partir do Ponto T de Goffman com Stopwords



Fonte: A autora (2014).

A partir do resultado percebe-se que há presença de stopwords é quase maior do que as palavras normais, e que somente por este motivo não seria um bom conteúdo

para uma indexação. De resto o resultado ficou semelhante aos demais. A imagem 23 apresenta as ementas com stopwords, desta vez com a fórmula alternativa:

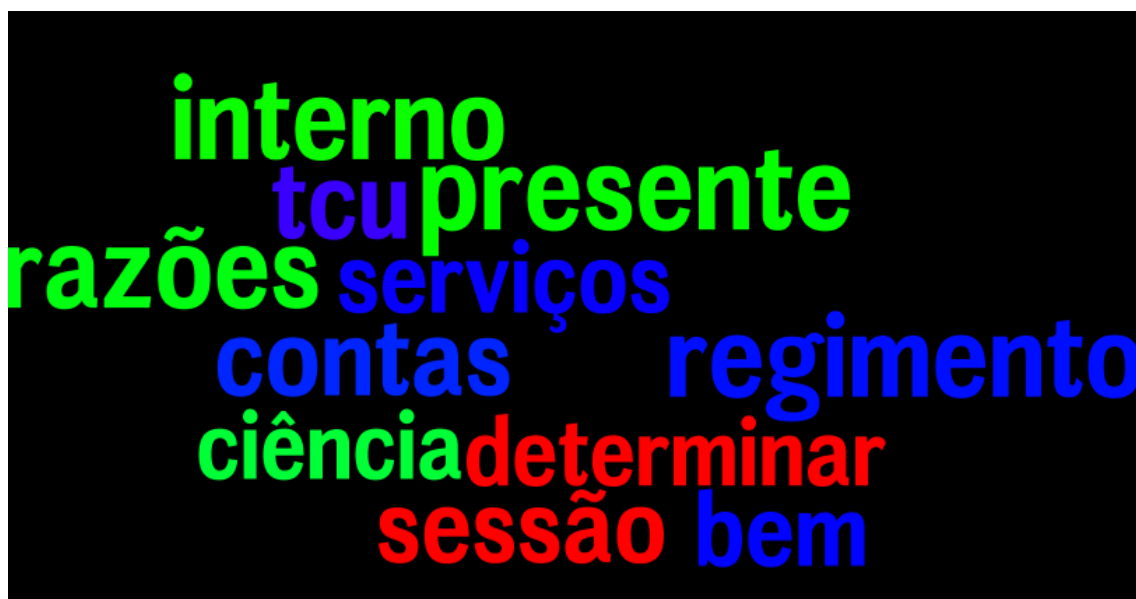
Imagem 23 – Análise na ementa a partir da Fórmula alternativa e com Stopwords



Fonte: A autora (2014).

O resultado acima, assim como o anterior possui uma quantidade muito grande de stopwords, impedindo a sua utilização para uma indexação ou recuperação possível da informação. Abaixo será analisado o campo acórdão, aplicando novamente as duas fórmulas. Conforme imagem 24:

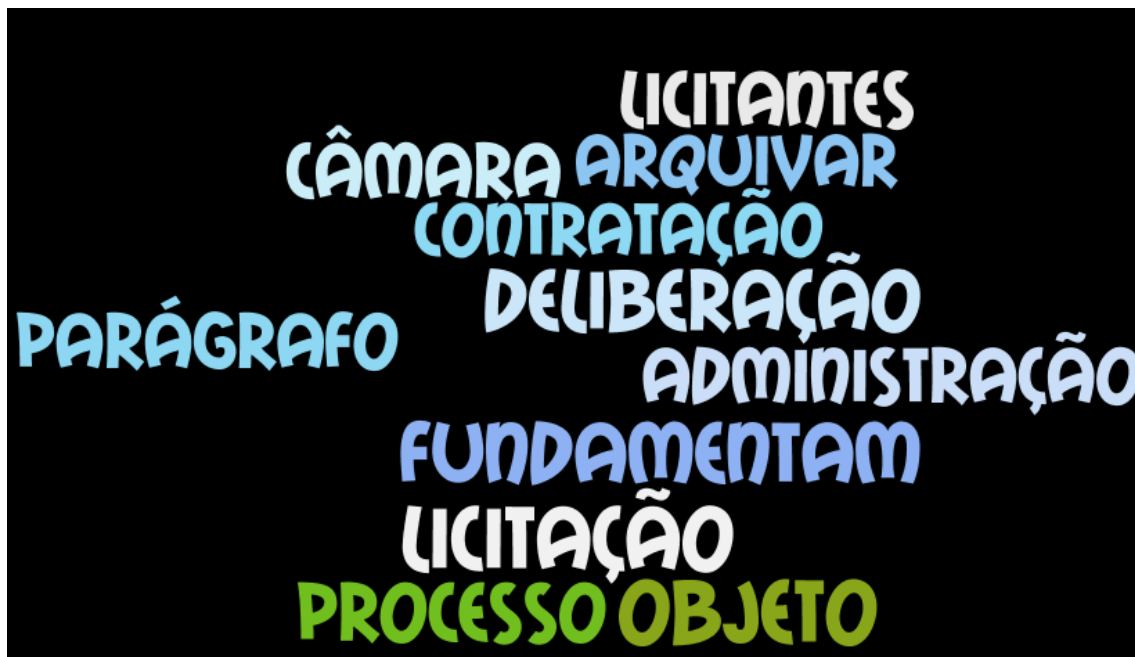
Imagem 24 – Análise no acórdão a partir do Ponto T de Goffman



Fonte: A autora (2014).

Como os outros o resultado não é satisfatório, em questão de assunto. Abaixo será testado na fórmula alternativa para comparação, conforme imagem 25:

Imagem 25 – Análise no acórdão a partir da fórmula alternativa



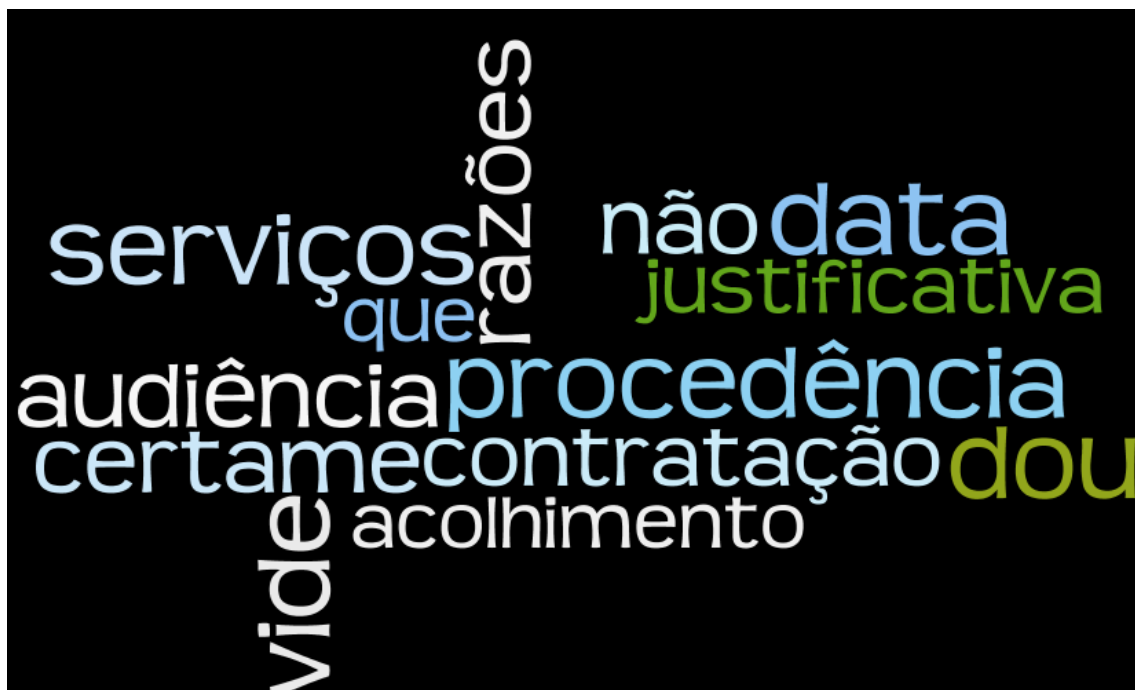
Fonte: A autora (2014).

O resultado acima, apesar de não possuir stopwords não apresenta relação com o assunto contido nos acórdãos. E, além disso, percebe-se que não possui muita semelhança com as palavras contidas no resultado anterior.

4.4.2.3 ANÁLISE EM NUVEM DE PALAVRAS PESQUISA LIVRE

Abaixo serão apresentados os resultados obtidos, referente às ementas dos acórdãos obtidos através da pesquisa livre. Abaixo segue imagem 26 do resultado obtido usando a fórmula do ponto T de Goffman:

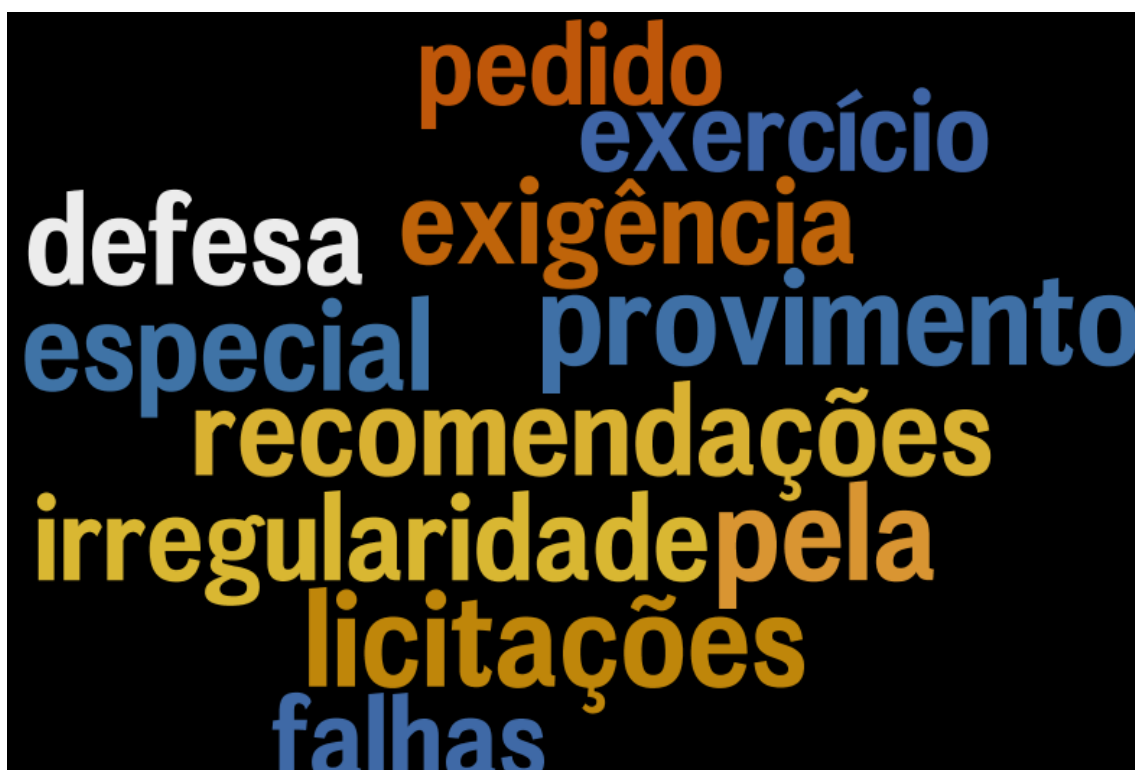
Imagem 26 – Análise das ementas (pesquisa livre) a partir do Ponto T de Goffman



Fonte: A autora (2014).

Pode-se observar através da figura que o resultado foi muito semelhante a nuvem de palavras gerada na pesquisa por assunto. E ainda que assim como na pesquisa por assunto o resultado foi insatisfatório, tanto referente ao possível assunto quanto a quantidade de stopwords obtidas. A seguir, na imagem 27, será demonstrado o resultado com a aplicação da fórmula alternativa:

Imagem 27 – Análise das ementas (pesquisa livre) a partir da fórmula alternativa



Fonte: A autora (2014).

Aqui observa-se que além de obter stopwords o resultado ainda apresenta muitas expressões que não fazem sentido junto ao contexto. Desta forma percebe-se que o resultado não atingiu as expectativas do presente trabalho.

4.4.2.4 ANÁLISE EM NUVEM DE PALAVRAS PESQUISA ESPECÍFICA

Neste item será apresentado a nuvem de palavras (imagem 28) dos resultados referente a pesquisa por assunto mais específico. Primeiramente a partir do ponto T de Goffman:

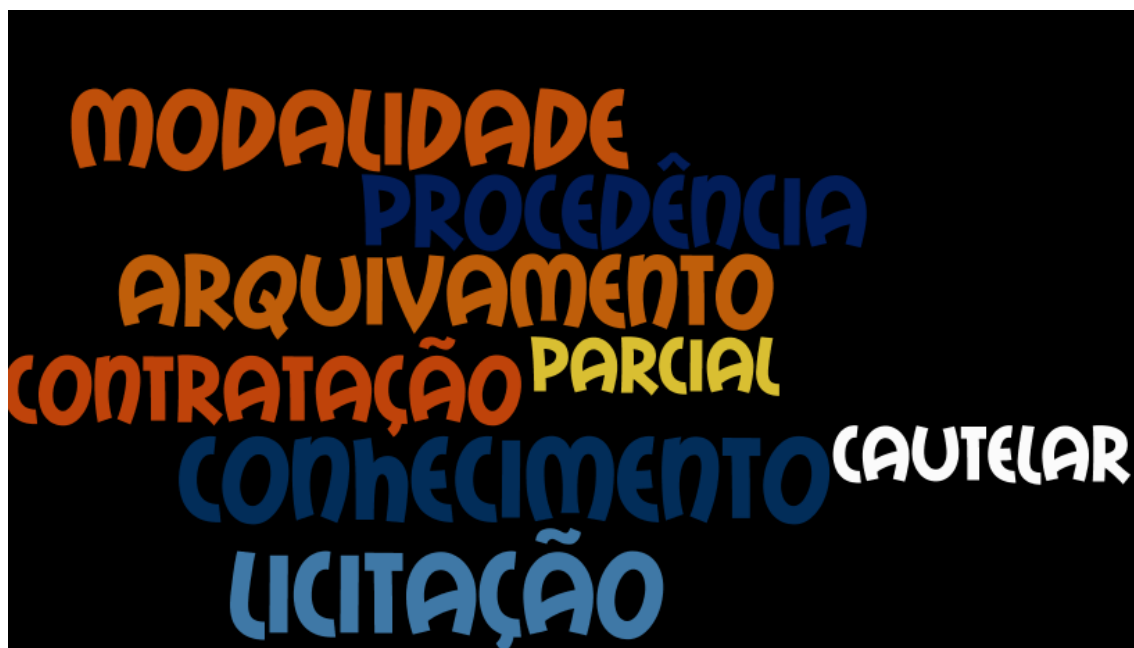
Imagem 28 – Nuvem de palavras pesquisa específica



Fonte: A autora (2014).

Percebe-se através do resultado que as palavras ficaram bem próximas de mostrar o assunto dos acórdãos escolhidos. Abaixo, na imagem 29, o resultado é apresentado através da fórmula alternativa:

Imagem 29 – Nuvem de palavras pesquisa específica 2



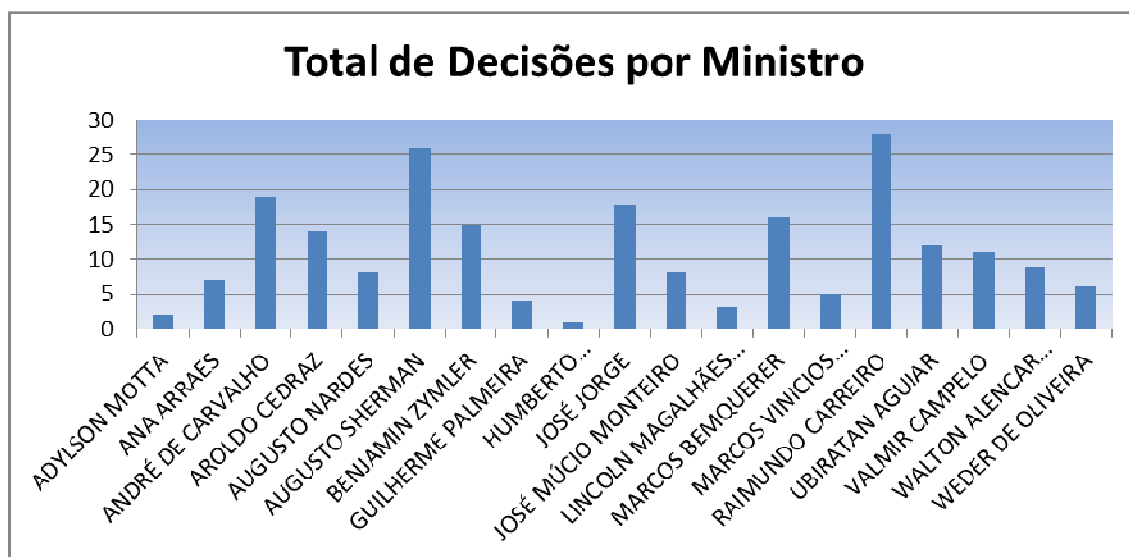
Fonte: A autora (2014).

Nestes os resultados ficaram bem semelhantes, mas pode-se perceber que de modo geral uma pesquisa em um assunto mais específico gera melhores resultados.

4.4.3 ANÁLISE DE AUTORIA

A seguir será apresentado e analisado o resultado referente aos dados obtidos referentes à autoria. A partir da tabulação dos dados, foi possível identificar o total de decisões de cada ministro e suas respectivas tendências de decisões, conforme gráfico 1 abaixo e apêndice B:

Gráfico 1 – Total de decisões por ministro



Fonte: A autora (2014).

Observa-se através do gráfico, que dos relatores mencionados, o que possui maior número de decisões é o ministro Raimundo Carreiro com 28 decisões, seguido do ministro Augusto Sherman com 26 decisões. Em contrapartida o ministro Humberto Guimarães Souto foi o que apareceu com um número menor de decisões com apenas uma decisão.

Pode-se ainda analisar a tendência de decisões de cada ministro, conforme exposto no apêndice A. Percebe-se que existe uma tendência maior por parte dos ministros para avaliar as decisões como procedentes, e ainda que o ministro Augusto Sherman tem uma tendência maior a acatar uma decisão como improcedente.

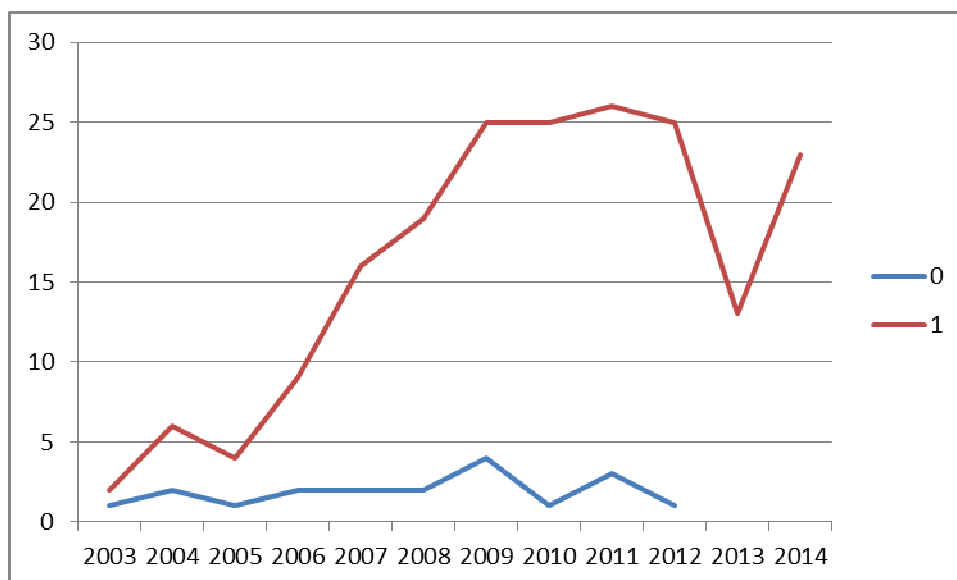
A partir dos dados é possível identificar fatos referente a autoria e tendências de decisões isoladas ou em conjunto, o que significa que o estudo atendeu neste sentido

seu propósito. Além disso, ainda é possível trabalhar esses dados com um volume maior de decisões para obter resultados mais completos.

4.4.4 ANÁLISE DA DECISÃO

Neste tópico serão apresentados os resultados obtidos referentes a fatores que envolvem possíveis padrões em decisões ao longo do tempo. Abaixo segue gráfico 2 que representa as decisões tomadas ao longo do tempo, onde 0 é igual a uma decisão improcedente e 1 é igual a uma decisão procedente:

Gráfico 2 – Decisões de 2003 a 2014



Fonte: A autora (2014).

Percebe-se que o maior número de decisões se concentrou entre os anos de 2008 e 2012 e em sua maior parte sendo procedentes, também pode-se observar que o maior número de decisões improcedentes ocorreu no ano de 2009. Assim como no ano de 2013 as decisões procedentes caíram muito em questão de quantidade, e as improcedentes tiveram resultado nulo. O mesmo se repetiu para as decisões improcedentes no ano de 2014, e para as procedentes o resultado foi reverso, aumentou significativamente novamente.

Percebe-se a partir dos dados que o volume de decisões subiu significativamente após 2008, o que caracteriza uma mudança no padrão de decisões ou até mesmo de produtividade dos ministros.

Esse tipo de análise é muito importante para descobrir mudanças ao longo do tempo e padrões. Assim como pode também mostrar algumas informações sobre produtividade como supõem-se no parágrafo anterior.

4.4.5 ANÁLISE DA CITAÇÃO – GENEALOGIA

Neste item será analisado o resultado referente aos dados que foram tabulados e manipulados em uma planilha eletrônica, referente à súmula 257/2010 do Tribunal de Contas da União, que trata do uso do pregão em contratações de serviços comuns de engenharia. Os dados serão analisados visando obter informações sobre sua origem e mostrando uma nova forma de tratar dados jurídicos, a análise é feita em suas referencias, as quais tratam-se de vários tipos de documentos jurídicos, como: acórdãos, leis, decretos, etc.

Os dados foram trabalhados a partir de análise de citações. Abaixo segue gráficos 3 e 4 com os resultados de citações feitas por tipo documental e em seguida por ano:

Gráfico 3 – Tipos Documentais mais citados



Fonte: A autora (2014).

Gráfico 4 – Citações por ano



Fonte: A autora (2014).

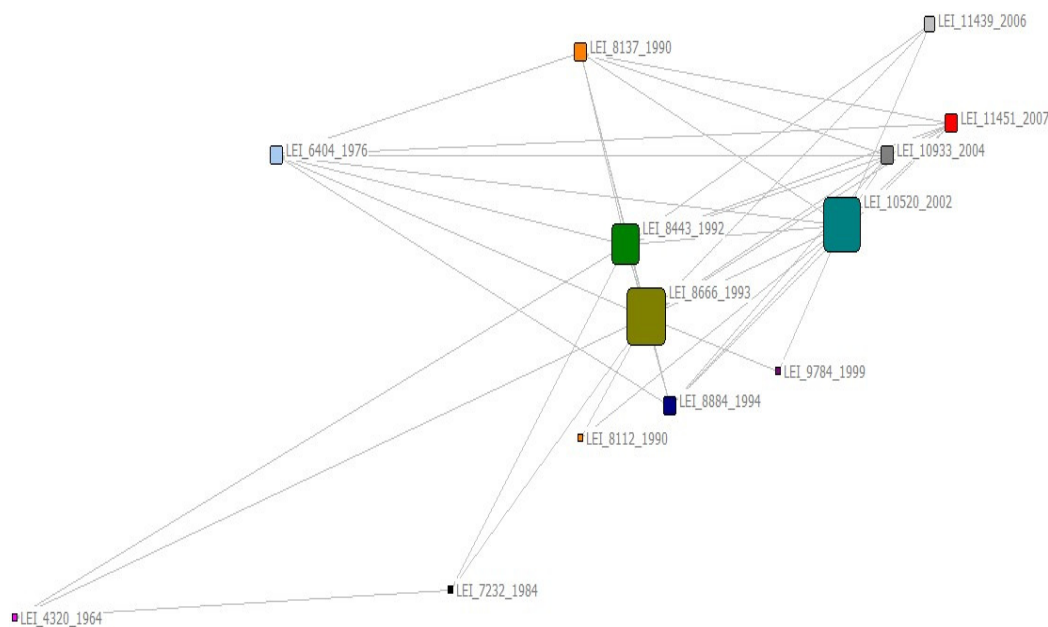
A partir dos dados pode-se afirmar que para composição da súmula em questão o acórdão foi o tipo documental mais citado pelos ministros relatores, com 46,15% de citações, e em contrapartida os Decreto leis foram os menos citados, juntamente com as súmulas, projetos de leis, instruções normativas e medidas provisórias, com 0,64% das citações.

Também é possível afirmar, que a partir do ano 2000 houve um período de “transição documental”, onde os acórdãos e jurisprudências ganharam um espaço maior de citações em decisões. O que explica o gráfico 1, que aponta os acórdãos como tipo documental mais citado.

É possível neste sentido, ainda analisar dados referentes a tipos documentais de forma isolada. Por exemplo, uma análise apenas de leis mais citadas.

Abaixo, na imagem 30, será apresentada uma rede de coocorrência. Pode-se observar que os nós maiores são os pontos mais centrais, ou seja, são as leis que mais receberam número de citação. Podemos presumir que seja por sua importância sobre determinado assunto.

Imagem 30 – Rede de Leis

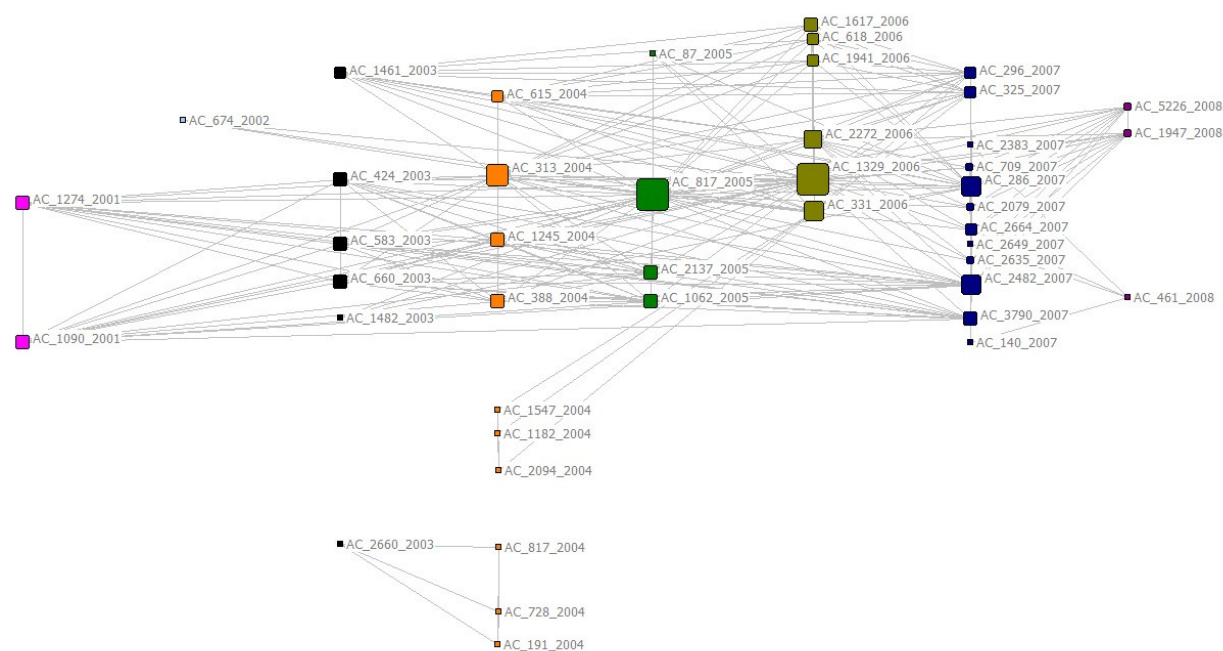


Fonte: A autora (2014).

Observa-se que, neste caso, os pontos centrais foram a lei 8.666/1993 (lei de licitações) e a lei 10.520/2002 (aborda questões referentes a modalidade pregão), o que confirma a teoria de que são os mais centrais por sua relevância dentro do assunto.

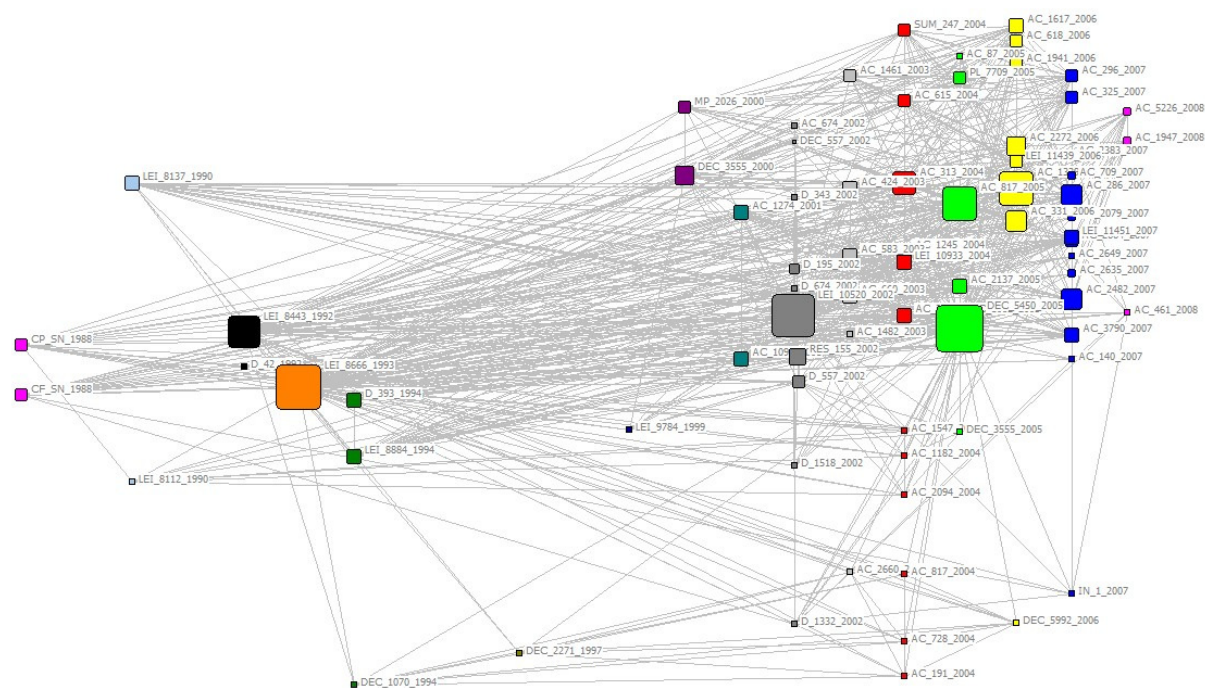
O estudo foi repetido, conforme ilustrado nas imagens 31 e 32, para os acórdãos mais citados, e em seguida para todos os tipos documentais em conjunto:

Imagem 31 – Acórdãos citados



Fonte: A autora (2014).

Imagem 32 – Todos os documentos citados



Fonte: A autora (2014).

A partir das imagens pode-se afirmar que os pontos mais centrais ou acórdãos mais citados foram os: AC 1329/2006 e 817/2005. Seguidos dos acórdãos 313/2004, 286/2007, 2482/2007, e todos são interligados entre si, ou seja, foram citados pelo menos uma vez em conjunto.

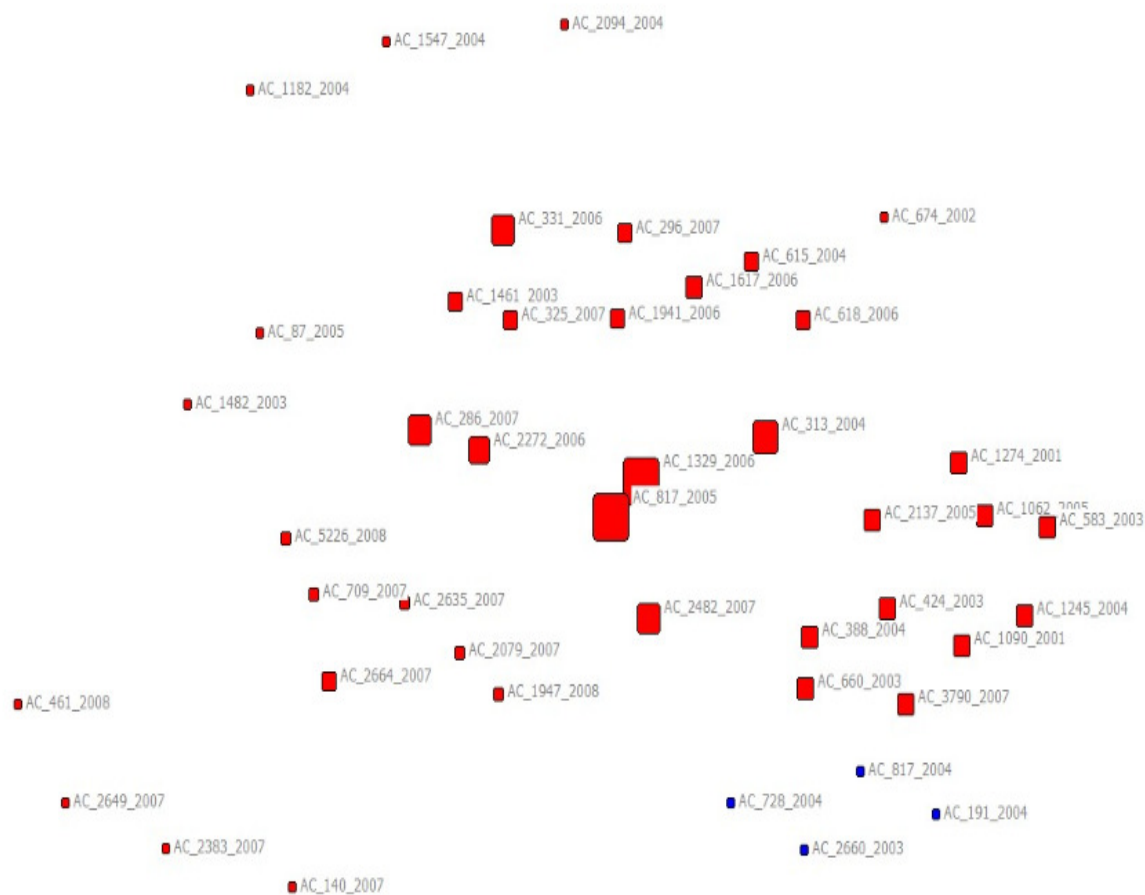
Já na segunda imagem, percebe-se que o resultado altera bastante quando inclui-se outros tipos documentais na mesma análise. Neste resultado a Lei 8.666/1993, o decreto 5450/2005, a lei 10.550/2002 e a lei 8.443/1992 são os mais citados, ou aparecem em uma quantidade maior de documentos.

Ainda na imagem um, por algum motivo, existem acórdãos que aparecem mais abaixo e sem conexão com os demais, pode-se afirmar que estes são acórdãos que foram citados em um ou mais documentos, mas sem qualquer relação com os demais, ou seja, quando são citados, não é citado nenhum dos outros acórdãos que aparecem acima. É possível supor que seja de assunto muito diferente dos demais.

A imagem 33 abaixo mostra a mesma figura de forma diferente, onde os anos não estão em ordem e onde fica bem mais clara essa “divisão” feita nos grupos de

acórdãos, além disso ela ainda nos dá algumas “pistas” de acórdãos mais citados na súmula.

Imagem 33 – Rede de coocorrência acórdãos



Fonte: A autora (2014).

A partir da análise da rede é possível afirmar que os acórdãos mais centrais são os que mais tem relação com o tema da súmula.

4.4.6 DISCUSSÃO SOBRE POSSIBILIDADES DE APLICAÇÃO

A proposta deste trabalho foi realizar alguns testes referentes à possibilidade de trabalhar com a informação jurídica de modo que ajude no tratamento e indexação desta informação em sites jurídicos, para facilitar a busca e recuperação da informação.

Para entender melhor essa proposta, supomos, por exemplo, que um usuário qualquer estivesse fazendo uma busca, assim como neste trabalho, por acórdãos com tema “pregão”. E assim que ele pesquisasse por esta palavra, ele obtive-se um resultado em forma de uma nuvem de palavras (para facilitar sua visualização), com diversos assuntos diferentes, dentro deste tema maior. A partir daí ele poderia escolher facilmente o assunto que estivesse mais próximo do que se deseja.

Após entrar no assunto, ele teria uma infinidade de acórdãos relacionados com o mesmo tema, seguindo uma linha no tempo, uma “árvore de opções” (conforme foi ilustrado na análise de citações). Onde seria possível observar os precedentes daquele acórdão, e seus acórdãos relacionados. Assim seria mais ágil o processo de busca por assuntos do mesmo assunto se o primeiro acórdão não o atende-se.

Outra forma de ajudar nesta busca e recuperação da informação seria inserir itens como análises de decisões e análises de autoria, por exemplo. Assim esses dois tópicos funcionariam como filtros de busca, onde o usuário poderia escolher na hora da busca, que tipo de decisão ele procura (procedente ou improcedente), qual o ano da decisão, qual ministro relator, entre outras.

Entende-se que esse trabalho seja apenas uma proposta para que a situação acima ocorrer um dia. No entanto, a partir das aplicações acima, foi possível visualizar que isso é possível. A informação jurídica pode ser tratada de modo que facilite a busca e recuperação da informação para usuários destes sites.

4.5 SÍNTESE DA SEÇÃO

A aplicação de técnicas infométricas nestes registros jurídicos permitiu mostrar que essa informação pode ser tratada e manipulada de modo quantitativo. Ainda que com certa dificuldade no que tange a organização de conteúdos e palavras-chaves.

O resultado referente à análise de palavras não foi satisfatório, possivelmente devido à união de assuntos diversos no mesmo tema pesquisado, o que nos faz pensar em uma falha na atual indexação ou talvez uma falta de refinamento da busca da informação.

Em contrapartida os resultados referentes à autoria, decisões e genealogia foram muito satisfatórios e sem qualquer prejuízo poderiam ser aprofundado no futuro. É possível organizar os dados e tratá-los de modo que se possa entender melhor o passado e até obter insights sobre o futuro da área jurídica.

A aplicação das técnicas infométricas nestes registros jurídicos permitiu a visualização de dados como:

- Tipos de decisões mais tomadas;
- Ministros mais produtivos;
- Mudanças em relação a volume de informações no decorrer dos anos;
- Leis, decretos, e outros tipos documentais mais recorridos ou citados;
- Volume de informação referente a uma área de conhecimento.

Pode-se então afirmar que além dessas áreas contribuírem muito uma com a outra, ainda há uma infinidade de possibilidades de pesquisa e possíveis contribuições maiores, já que o estudo foi com um volume relativamente pequeno de registros jurídicos.

5 CONSIDERAÇÕES FINAIS

Esta seção sintetiza o estudo realizado a partir da retomada dos objetivos específicos propostos e a resposta à pergunta que norteou o trabalho, levantando as principais contribuições, bem como suas limitações e sugestões para pesquisas futuras.

Inicialmente, se faz necessário resgatar os objetivos específicos propostos:

- Verificar os procedimentos de busca de jurisprudência num site jurídico, representado no presente estudo pelo site do Tribunal de Contas da União: Conforme descrito na Seção 4.1 do presente trabalho foi realizada a análise no site do Tribunal de Contas da União. O resultado da pesquisa aponta uma dificuldade de busca e recuperação das informações, ainda que apresentem várias formas de pesquisa.
- Medir registros jurídicos enquanto unidade informacional: Conforme descrito na Seção 4.2, foi feito um estudo do acórdão como unidade de informação, dividindo-o em metadados. Este estudo permitiu demonstrar que esse tipo de registro pode ser particionado para um estudo quantitativo da informação.
- Medir registros jurídicos selecionados a partir de técnicas infométricas: Na Seção 4.3 foram aplicadas diferentes técnicas informétricas, como nuvens de palavras, redes de cooperação, Ponto T de Goffman, e ainda foram feitos testes com tabelas dinâmicas. Os resultados foram interpretados e a partir disso foi possível responder a problemática do trabalho.
- Verificar a possibilidade de extensão dos resultados obtidos em processos de busca e recuperação da informação: Na seção 4.6 foi discutido a possibilidade do estudo futuramente ajudar na busca e recuperação da informação, considerando que ainda há estudo para se fazer referente ao assunto. Essa discussão permitiu demonstrar que a proposta do trabalho pode ser trabalhada e implementada.

Assim é possível afirmar que é possível sim tratar e visualizar melhor a informação disposta. Através do estudo pode-se perceber que além da informação oferecer subsídios para ser trabalhada quantitativamente, ainda há várias técnicas que podem ser usadas de diferentes formas para trabalhar esses registros, como: nuvem de palavras, redes de ocorrência, fórmula de Goffman, análise e manipulação em planilhas

eletrônicas. Além disso, é possível trabalhar com diferentes aspectos desses registros, como: conteúdo (palavras), decisões, autorias, dados temporais, entre outros.

5.1 LIMITAÇÕES E SUGESTÕES PARA PESQUISAS FUTURAS

O estudo apresentou algumas limitações que serão descritas e a partir destas serão apresentadas propostas para pesquisas futuras. Neste estudo não foi possível selecionar um número maior de registros jurídicos, devido ao tempo disposto. Recomenda-se para as próximas pesquisas ampliar esse número e também refinar mais o processo de busca para assuntos bem específicos, para verificar se os resultados referentes à análise de palavras alterem em algum sentido e seja mais satisfatório.

Ainda para próximas pesquisas recomenda-se tratar os resultados de uma forma qualitativa, verificando também as necessidades dos usuários que utilizam esses sites e buscam essas informações diariamente. Bem como, procurar buscar a opinião e consultoria referente a resultados obtidos, para verificar se eles podem ser utilizado ou se alguma forma faz algum sentido.

5.2 COMENTÁRIOS FINAIS

Esse estudo permitiu verificar o processo de busca de informações em sites jurídicos, assim como medir a dificuldade do usuário no que diz respeito a principalmente a recuperação desta informação.

Além disso, permitiu demonstrar que é possível trabalhar com a informação jurídica de forma quantitativa, a fim de melhorar sua visualização e consequentemente o entendimento de que a utiliza em seu dia-a-dia.

Isso pode ocorrer por várias técnicas como rede de coocorrência, nuvens de palavras e outras formas mais visuais que ajudem o usuário em sua procura.

É possível também afirmar que a infometria pode ter uma contribuição significativa em outras áreas do conhecimento além da ciência da informação. Ampliando seu campo de estudo para áreas diversas.

REFERÊNCIAS

BUFREM, L. PRATES, Y. **O saber científico registrado e as práticas de mensuração da informação.** Ciência da Informação, Brasília, v. 34, n. 2, p. 9-25, 2005. Disponível em: < <http://revista.ibict.br/ciinf/index.php/ciinf/article/viewFile/682/587>>. Acesso em: 06/05/2014.

GUEDES, V. L. S. **A Bibliometria e a Gestão da Informação e do conhecimento científico e tecnológico: uma revisão da literatura.** Ponto de Acesso, Salvador, V.6, n. 2, p. 74-109 ago 2012. Disponível em: < <file:///C:/Documents%20and%20Settings/suzana.souza/Meus%20documentos/Downloads/5695-19228-2-PB.pdf> >. Acesso em: 10/10/2014.

MACIAS-CHAPULA, Cesar A. **O papel da informetria e da cienciometria e sua perspectiva nacional e internacional.** Ciência da Informação, Brasília, v. 27, n. 2, p. 134-140, maio/ago. 1998. Disponível em: < <http://www.scielo.br/pdf/ci/v27n2/macias.pdf> >. Acesso em 05/05/2014.

MAXIMILIANO, C. **Hermenêutica e aplicação do Direito.** 20 ed. – Rio de Janeiro: Forense, 2011.

MIRANDA, A. C. C. D'AMORE, T. M. PINTO, V. B. **Gestão documental da informação jurídica.** Perspectivas em Ciência da Informação, v.18, n.3, p.96-110, jul./set. 2013.

MOTA, F. R. L. **Registro de informação no sistema de informação em saúde: um estudo das bases SINASC, SIAB e SIM, no estado de Alagoas.** 2009. 266 f. Tese (Doutorado em Ciência da Informação) – Universidade Federal de Minas Gerais, Minas Gerais. 2009. Disponível em: <<http://dx.doi.org/10.1590/S1413-99362009000300020>>. Acesso em: 05/05/2014.

SANTIN, D. M. **Avanços e perspectivas da infometria e dos indicadores multidimensionais na análise de fluxos da informação e estruturas do conhecimento.** Revista Eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, v. 16, n. 32, p. 107-122, 2011. Disponível em: < <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2011v16n32p107/19343> >. Acesso em: 06/05/2014.

SANTOS, R. N. M; KOBASHI, N. Y. **Bibliometria, cientometria, infometria: conceitos e aplicações.** Tendências da Pesquisa Brasileira em Ciência da Informação, Brasília, v. 2, n. 1, p. 155-172, jan./dez. 2009. Disponível em: <<http://inseer.ibict.br/ancib/index.php/tpbci/article/view/21/43>>. Acesso em: 07/05/2014.

SPINAK, Ernesto. **Diccionario enciclopédico de bibliometría, cienciometría e informetria.** Caracas: Cresalc/Unesco, 1996.

SHINTAKU, M. ROBREDO, J. BAPTISTA, D. C. **Webometria dos Repositórios institucionais acadêmicos.** Ciência da Informação, Brasília, v. 40 n. 2, p. 312-326, maio/ago., 2011.

SOUZA, T. B. de. CATARINO, M. E. SANTOS, P. C. dos. **Metadados: Catalogando dados na internet**. Transinformação, Amapá, v. 9, n. 2, p. 93-105, maio/ago., 1997.

SOUZA, M. I. F. VENDRUSCULO, L. G. MELO, G. C. **Metadados para a descrição de recursos de informação eletrônica: utilização do padrão Dublin Core**. Ciência da Informação, Brasília, v. 29, n. 1, p. 93-102, jan./abr. 2000.

SANTOS, E. L; COELHO, M. F. P. **Portal de periódicos da CAPES: um estudo da sua utilização pelos usuários da faculdade de Direito da Universidade Federal de Minas Gerais**. In: . Bibliometria e cientometria: estudos temáticos. São Carlos: Pedro e João Editoras. 2013. cap. p.319-320.

APÊNDICE A – TABELA DE DECISÕES DE CADA MINISTRO

1 = Decisão Procedente

0 = Decisão Improcedente

Relatores	0	1	Total Geral
ADYLSO MOTA	1	1	2
2004	1	1	2
ANA ARRAES		7	7
2012		3	3
2013		1	1
2014		3	3
ANDRÉ DE CARVALHO	1	18	19
2008		1	1
2009		5	5
2010		2	2
2011		4	4
2012	1	2	3
2013		1	1
2014		3	3
AROLD DO EDRAZ	1	13	14
2007		1	1
2008		2	2
2010		2	2
2011	1	4	5
2013		1	1
2014		3	3
AUGUSTO NARDES	2	6	8
2006	1		1
2008		2	2
2009	1		1
2011		2	2
2012		2	2
AUGUSTO SHERMAN	4	22	26
2003		1	1
2007	1		1
2009	1	5	6
2010	1	3	4
2011	1	4	5
2012		2	2
2013		3	3
2014		4	4
BENJAMIN ZYMLER	1	14	15
2006		2	2
2007		3	3
2008		4	4
2009	1	2	3
2010		1	1

2013		1	1
2014		1	1
GUILHERME PALMEIRA	1	3	4
2006		1	1
2007	1	1	2
2008		1	1
HUMBERTO GUIMARÃES SOUTO	1		1
2003	1		1
JOSÉ JORGE		18	18
2009		2	2
2010		4	4
2011		5	5
2012		3	3
2013		3	3
2014		1	1
JOSÉ MÚCIO MONTEIRO		8	8
2010		2	2
2011		1	1
2012		4	4
2014		1	1
LINCOLN MAGALHÃES DA ROCHA	1	2	3
2005	1	2	3
MARCOS BEMQUERER		16	16
2004		1	1
2006		1	1
2007		3	3
2008		1	1
2009		4	4
2010		1	1
2011		3	3
2014		2	2
MARCOS VINÍCIOS VILAÇA		5	5
2005		1	1
2006		1	1
2007		2	2
2008		1	1
RAIMUNDO CARREIRO	1	27	28
2007		3	3
2008		5	5
2009	1	4	5
2010		6	6
2012		4	4
2013		2	2
2014		3	3
UBIRATAN AGUIAR	2	10	12
2004		3	3
2006		3	3
2007		2	2
2008	2	2	4
VALMIR CAMPELO	2	9	11
2006	1		1
2007		1	1
2009		2	2
2010		1	1

2011	1		1
2012		4	4
2013		1	1
WALTON ALENCAR RODRIGUES	1	8	9
2003		1	1
2004	1	1	2
2005		1	1
2006		1	1
2009		1	1
2010		1	1
2011		1	1
2012		1	1
WEDER DE OLIVEIRA		6	6
2010		2	2
2011		2	2
2014		2	2
<hr/>			
Total Geral	19	193	212

**APÊNDICE B – TABELA DE CONTAGEM DE NÚMERO DE VEZES
QUE A PALAVRA “PREGÃO” APARECE EM CADA DOCUMENTO
(PESQUISA POR ASSUNTO)**

Nº do resultado/linha de resultado	Nº do acórdão	Sumário	Relatório do ministro relator	Voto	Acórdão
01/1	<u>AC-1542-21/14-P</u>	1	0	1	1
02/1	<u>AC-1929-27/13-P</u>	1	39	3	3
03/1	<u>AC-1615-23/13-P</u>	1	17	8	2
04/1	<u>AC-0959-13/13-P</u>	0	41	1	2
05/1	<u>AC-2766-15/12-1</u>	1	16	3	2
06/1	<u>AC-0518-07/12-P</u>	1	14	3	2
07/1	<u>AC-0169-03/12-P</u>	1	20	3	2
08/1	<u>AC-0478-02/11-1</u>	1	72	0	0
09/1	<u>AC-0339-06/10-P</u>	1	41	4	7
10/1	<u>AC-5804-36/09-1</u>	1	2	0	0
11/1	<u>AC-2143-37/09-P</u>	1	14	11	0
12/1	<u>AC-0694-09/14-P</u>	1	25	23	3
13/1	<u>AC-2627-37/13-P</u>	0	19	2	4
14/1	<u>AC-2883-41/13-P</u>	1	11	2	1
15/1	<u>AC-6240-38/13-2</u>	1	4	0	1

01/2	<u>AC-2166-32/14-P</u>	0	12	6	2
02/2	<u>AC-0392-05/11-P</u>	1	18	8	4
03/2	<u>AC-1513-22/13-P</u>	1	19	5	3
04/2	<u>AC-1153-16/13-P</u>	0	16	8	2
05/2	<u>AC-1888-28/10-P</u>	0	28	17	1
06/2	<u>AC-0718-04/10-1</u>	0	9	4	0
07/2	<u>AC-2183-40/08-P</u>	2	36	8	4
08/2	<u>AC-2816-50/09-P</u>	0	12	7	1
09/2	<u>AC-1946-12/10-1</u>	0	3	0	0
01/3	<u>AC-1188-16/11-P</u>	1	30	10	4
02/3	<u>AC-1462-22/10-P</u>	1	23	5	5
03/3	<u>AC-0694-09/14-P</u>	1	21	0	0
04/3	<u>AC-0122-02/12-P</u>	1	27	0	0
05/3	<u>AC-1339-19/11-P</u>	1	34	0	0
06/3	<u>AC-0536-07/11-P</u>	1	35	1	2
01/4	<u>AC-1401-10/14-2</u>	0	0	2	0
02/4	<u>AC-0694-09/14-P</u>	1	25	0	0
03/4	<u>AC-0620-08/14-P</u>	1	27	0	0
04/4	<u>AC-2074-33/11-P</u>	1	11	7	4
05/4	<u>AC-2077-33/11-P</u>	2	34	0	0
01/5	<u>AC-1168-20/09-P</u>	2	52	30	2
01/6	<u>AC-7724-33/11-2</u>	1	45	11	0

02/6	<u>AC-0694-09/14-P</u>	1	25	23	3
03/6	<u>AC-3368-48/13-P</u>	0	33	1	0
04/6	<u>AC-10604-39/11-2</u>	1	21	5	3

**APÊNDICE C – TABELA DE CONTAGEM DE NÚMERO DE VEZES
QUE A PALAVRA “PREGÃO” APARECE EM CADA DOCUMENTO
(PESQUISA LIVRE)**

Número do resultado	Número do Acórdão	Sumário	Relatório	Voto	Acórdão
01	<u>AC-2073-29/04-1</u>	0	2	0	0
02	<u>AC-0601-08/07-1</u>	2	15	0	0
03	<u>AC-1068-12/07-1</u>	0	4	0	0
04	<u>AC-1305-14/07-1</u>	0	4	10	0
05	<u>AC-2336-26/08-1</u>	0	1	0	0
06	<u>AC-2105-13/09-1</u>	0	1	0	0
07	<u>AC-2712-16/09-1</u>	0	2	0	0
08	<u>AC-5365-33/09-1</u>	0	1	0	0
09	<u>AC-6483-40/09-1</u>	0	1	0	0
10	<u>AC-3146-18/10-1</u>	0	0	1	0
11	<u>AC-3143-18/10-1</u>	0	8	0	0
12	<u>AC-3319-19/10-1</u>	0	5	0	0
13	<u>AC-3493-20/10-1</u>	0	3	4	0
14	<u>AC-5341-23/11-1</u>	0	6	0	0
15	<u>AC-8673-35/11-1</u>	0	12	3	1
16	<u>AC-9859-41/11-1</u>	0	10	2	0
17	<u>AC-5613-33/12-1</u>	1	9	4	2

18	<u>AC-7498-44/12-1</u>	0	1	0	0
-----------	--	---	---	---	---

Agora será tabulado o termo de busca sistema “S” para comparar os resultados obtidos:

Número do resultado	Número do Acórdão	Sumário	Relatório	Voto	Acórdão
01	<u>AC-2073-29/04-1</u>	0	1	1	0
02	<u>AC-0601-08/07-1</u>	0	6	0	0
03	<u>AC-1068-12/07-1</u>	0	1	1	0
04	<u>AC-1305-14/07-1</u>	0	1	1	0
05	<u>AC-2336-26/08-1</u>	2	6	0	0
06	<u>AC-2105-13/09-1</u>	0	7	3	0
07	<u>AC-2712-16/09-1</u>	0	4	1	0
08	<u>AC-5365-33/09-1</u>	0	4	0	0
09	<u>AC-6483-40/09-1</u>	0	2	0	0
10	<u>AC-3146-18/10-1</u>	0	8	7	0

11	<u>AC-3143-18/10-1</u>	0	12	0	0
12	<u>AC-3319-19/10-1</u>	0	10	0	0
13	<u>AC-3493-20/10-1</u>	0	6	0	0
14	<u>AC-5341-23/11-1</u>	0	12	1	0
15	<u>AC-8673-35/11-1</u>	0	2	1	0
16	<u>AC-9859-41/11-1</u>	0	4	3	0
17	<u>AC-5613-33/12-1</u>	0	3	1	0
18	<u>AC-7498-44/12-1</u>	0	2	0	0